

**UNIVERZITA KARLOVA**  
**PŘÍRODOVĚDECKÁ FAKULTA**  
**Katedra demografie a geodemografie**

Studijní program: Demografie

Studijní obor: Demografie



**Bc. Vojtěch Říha**

Vybrané metodické přístupy k tvorbě regionální populační prognózy:  
případová studie na úrovni Jihočeského kraje

Selected Methodical Approaches to Regional Population Forecast:  
A case study in the South Bohemian Region

Diplomová práce

Vedoucí závěrečné práce: RNDr. Klára Hulíková Tesárková, Ph.D.

Praha, 2017

**Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 20. 6. 2017

.....

podpis

## **Poděkování**

Rád bych na tomto místě poděkoval RNDr. Kláře Hulíkové Tesárkové, Ph.D. za vedení mé diplomové práce a za podnětné návrhy, které ji obohatily. Děkuji RNDr. Borisi Burcinovi, Ph.D. za poskytnutí projekce Jihočeského kraje pro závěrečnou komparaci výsledků a za odborné konzultace k diplomové práci. Rád bych také poděkoval doc. RNDr. Karlu Zvárovi, CSc. za poskytnutí konzultace k diplomové práci. Mé poděkování rovněž patří Českému statistickému úřadu, jmenovitě Mgr. Radku Havlovi za poskytnutí dat. Zároveň bych chtěl poděkovat své rodině za podporu během celého mého studia.

## **Vybrané metodické přístupy k tvorbě regionální populační prognózy: případová studie na úrovni Jihočeského kraje**

### **Abstrakt**

Cílem této práce je představit vybrané metodické přístupy k populačním prognózám, se zaměřením na regionální úroveň a s uvažováním různé délky časové řady. Konkrétní postupy jsou prakticky aplikovány pro potřeby vytvoření prognózy Jihočeského kraje. V teoretické části práce jsou stanoveny fáze zpracování populačních prognóz. Charakterizována je kohortně komponentní metoda s migrací, která se používá k vytvoření populačních prognóz. Dále jsou popsány vybrané analytické modely a funkce pro tvorbu dílčích prognóz úmrtnosti, plodnosti a migrace včetně nepřímé metody výpočtu migračního salda. Pro extrapolaci parametrů se v analýze časových řad charakterizují vybrané trendové funkce a Box-Jenkinsova metodologie. Analytická část práce se zaměřuje na zpracování prognózy Jihočeského kraje z krátké výchozí časové řady a z dlouhé výchozí časové řady. Z krátké výchozí časové řady je nejprve zpracována dílčí prognóza úmrtnosti analyzovaná Heligman-Pollardovým modelem, dále dílčí prognóza plodnosti analyzovaná Beta funkcí a dílčí prognóza migrace analyzovaná 25%, 50% a 75% kvantilem migračního salda v čase podle jednotek věku. Pro zjištění vzájemných vztahů mezi parametry Heligmana-Pollarda a Beta funkce je využita faktorová analýza. Z faktorové analýzy se získají faktorové skóre, která jsou extrapolována prostřednictvím trendových funkcí. Z těchto dílčích prognóz je poté vytvořena prognóza Jihočeského kraje do 1. 1. 2025. Dílčí prognóza úmrtnosti, plodnosti, imigrace a emigrace z dlouhé výchozí časové řady je zpracována prostřednictvím Lee-Carterových modelů. Index úrovně úmrtnosti a plodnosti je extrapolován Box-Jenkinsovou metodologií. Pro extrapolaci indexu úrovně emigrace a imigrace byly zvoleny trendové funkce. Prostřednictvím kohortně komponentní metody s migrací je poté vytvořena z dlouhé výchozí časové řady plnohodnotná populační prognóza Jihočeského kraje do 31. 12. 2035, která je na závěr porovnána se střední variantou prognózy vytvořenou v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu.

**Klíčová slova:** kohortně komponentní metoda s migrací, Lee-Carterův model, Heligman-Pollardův model, Beta funkce, nepřímá metoda výpočtu migračního salda

## **Selected Methodical Approaches to Regional Population Forecast: A case study in the South Bohemian Region**

### **Abstract**

The aim of this thesis is to introduce selected methodological approaches to population forecasts, focusing on the regional level and considering different lengths of time series. Specific procedures are applied to create a population forecast for the South Bohemian Region. In the theoretical part of this thesis, the stages of population forecasts processing are determined. The Cohort Component method with migration, which can be used to create population forecast, is characterized. Another part describes selected analytical models and functions for partial mortality, fertility and migration forecasts, including Indirect estimation of net migration. To extrapolate parameters, selected trending functions and the Box-Jenkins methodology are characterized in the part of the time series analysis. The analytical part of this thesis focuses on the creation of the South Bohemian Region forecast from short initial time series and long initial time series. From short initial time series, the partial forecast of mortality is analyzed by the Heligman-Pollard model, the partial forecast of fertility is analyzed by the Beta function and the partial forecast of migration is analyzed by 25%, 50% and 75% quartiles of the migration balance in time by age. To analyze a relationship between Heligman-Pollard and Beta function parameters, a factor analysis is used. From the factor analysis, factor scores are obtained that are extrapolated by trending functions. From these partial forecasts the forecast of the South Bohemian Region until 1. 1. 2025 is created. The partial forecasts of mortality, fertility, immigration and emigration from the long initial time series are processed by Lee-Carter's models. The Mortality and Fertility Index are extrapolated by the Box-Jenkins Methodology. For extrapolating the emigration and immigration level index, trending functions were selected. By Cohort Component method with migration is then created a full-value population forecast of the South Bohemian Region until 31. 12. 2035, which is compared with the middle version of the forecast created by the Water Research Institute T. G. Masaryk and with the projection of the Czech Statistical Office.

**Keywords:** Cohort Component Method with migration, Lee-Carter model, Heligman-Pollard model, Beta function, Indirect estimation of net migration

## OBSAH

<b>Seznam tabulek .....</b>	<b>8</b>
<b>Seznam obrázků .....</b>	<b>10</b>
<b>Seznam použitých zkratk.....</b>	<b>13</b>
<b>1 Úvod .....</b>	<b>14</b>
1.1 Cíle práce.....	15
1.2 Struktura práce .....	16
<b>2 Vymezení populačních prognóz.....</b>	<b>17</b>
2.1 Základní pojmy charakterizující populační prognózy .....	17
2.2 Obecná formulace prognózování a stanovení fází zpracování populačních prognóz.....	18
<b>3 Vybrané metody tvorby populačních prognóz včetně vybraných modelů směřujících k vytvoření dílčí prognózy úmrtnosti, plodnosti a migrace .....</b>	<b>21</b>
3.1 Kohortně komponentní metoda .....	22
3.1.1. Kohortně komponentní metoda bez migrace .....	23
3.1.2. Kohortně komponentní metoda s migrací .....	32
3.2 Vybrané modely při studiu očekávaného vývoje úmrtnosti .....	38
3.3 Vybrané modely při studiu očekávaného vývoje plodnosti.....	42
3.4 Přístup k analýze migrace.....	46
<b>4 Vybrané přístupy k analýze časových řad s potenciálním využitím pro demografické prognózování.....</b>	<b>51</b>
4.1 Vybrané trendové funkce .....	52
4.2 Box-Jenkinsova metodologie .....	53
4.2.1. Vybrané stacionární procesy .....	58
4.2.2. Nestacionární proces ARIMA.....	59
<b>5 Východiska pro zpracování regionální prognózy obyvatel Jihočeského kraje .....</b>	<b>62</b>
5.1 Legislativní změny ve vymezení území Jihočeského kraje.....	62
5.2 Datové zdroje využití v analytické části práce a jejich přímá souvislost s prahem a horizontem prognózy .....	63
<b>6 Prognóza vývoje počtu obyvatel Jihočeského kraje prostřednictvím analyzování vývoje krátké výchozí časové řady .....</b>	<b>65</b>
6.1 Dílčí prognóza úmrtnosti s využitím modelu Heligmana-Pollarda .....	66
6.2 Dílčí prognóza plodnosti s využitím Beta funkce .....	72
6.3 Výpočet migračního salda prostřednictvím nepřímé metody včetně prognózy jeho očekávaného vývoje .....	75

6.4	Populační prognóza počtu a věkové struktury mužů a žen v Jihočeském kraji zpracovaná z dílčích prognóz úmrtnosti, plodnosti a migrace analýzou krátké výchozí časové řady .....	77
<b>7</b>	<b>Prognóza vývoje počtu obyvatel Jihočeského kraje prostřednictvím analyzování vývoje dlouhé výchozí časové řady.....</b>	<b>81</b>
7.1	Dílčí prognóza úmrtnosti zpracovaná s využitím Lee-Carterova modelu úmrtnosti.....	82
7.2	Dílčí prognóza plodnosti zpracovaná s využitím Lee-Carterova modelu plodnosti.....	90
7.3	Dílčí prognóza migrace .....	94
7.3.1.	Dílčí prognóza imigrace s využitím Lee-Carterova modelu imigrace .....	94
7.3.2.	Dílčí prognóza emigrace s využitím Lee-Carterova modelu emigrace .....	98
7.3.3.	Výsledné migrační saldo mužů a žen .....	102
7.4	Populační prognóza počtu a věkové struktury mužů a žen v Jihočeském kraji zpracovaná z dílčích prognóz úmrtnosti, plodnosti a migrace prostřednictvím Lee-Carterových modelů.....	103
<b>8</b>	<b>Komparace výsledků prognózy Jihočeského kraje s 95% spolehlivostí vytvořené z dlouhé výchozí časové řady se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu .....</b>	<b>107</b>
8.1	Porovnání dílčích prognóz úmrtnosti, plodnosti a migrace se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu .....	107
8.2	Porovnání očekávaného vývoje věkové struktury obyvatel Jihočeského kraje se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu .....	110
<b>9</b>	<b>Závěr.....</b>	<b>114</b>
	<b>Seznam použité literatury.....</b>	<b>117</b>
	<b>Seznam použitých elektronických zdrojů .....</b>	<b>124</b>
	<b>Seznam použitých softwarových produktů.....</b>	<b>127</b>
	<b>Seznam použitých datových zdrojů.....</b>	<b>128</b>
	<b>Přílohy.....</b>	<b>134</b>

## SEZNAM TABULEK

Tab. 1	Seznam symbolů použitých u kohortně komponentní metody bez migrace.....	25
Tab. 2	Seznam symbolů použitých u kohortně komponentní metody s migrací .....	33
Tab. 3	Seznam symbolů použitých u nepřímé metody výpočtu migračního salda .....	47
Tab. 4	Seznam symbolů použitých u Box-Jenkinsovy metody .....	57
Tab. 5	Výsledek faktorové analýzy při hodnocení vztahů parametrů Heligmana-Pollarda u mužů .....	67
Tab. 6	Výsledek faktorové analýzy při hodnocení vztahů parametrů Heligmana-Pollarda u žen.....	68
Tab. 7	Odhad logaritmického trendu faktorových skóre jednotlivých faktorů Heligmana-Pollarda mužů a žen .....	68
Tab. 8	Vlastní čísla z faktorové analýzy při hodnocení počtu nezávislých parametrů Beta funkce .....	72
Tab. 9	Výsledek faktorové analýzy při hodnocení vztahů parametrů Beta funkce .....	73
Tab. 10	Očekávané příspěvky ke změně naděje dožití při narození mezi lety 2015–2035 ve věkových skupinách 0–14, 15–64 a 65 a více let .....	88
Tab. 11	Prognóza naděje dožití při narození chlapců a dívek v Jihočeském kraji se spolehlivostí 95 % zpracovaná z dlouhé výchozí časové řady včetně střední varianty prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce Českého statistického úřadu mezi lety 2016–2035 .....	108
Tab. 12	Prognóza úhrnné plodnosti Jihočeského kraje se spolehlivostí 95 % zpracovaná z dlouhé výchozí časové řady včetně střední varianty prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce Českého statistického úřadu mezi lety 2016–2035 .....	109
Tab. 13	Prognóza celkového migračního salda obyvatel Jihočeského kraje se spolehlivostí 95 % zpracovaná z dlouhé výchozí časové řady včetně střední varianty prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce Českého statistického úřadu mezi lety 2016–2035, vývoj celkového migračního salda mezi lety 1993–2015 .....	110
Tab. 14	Střední varianta prognózy věkové struktury mužů a žen v Jihočeském kraji k 1. 1. vytvořená v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce věkové struktury mužů a žen v Jihočeském kraji k 1. 1. zpracovaná Českým statistickým úřadem, roky 2017, 2026 a 2036.....	111



---

Tab. 15	Prognóza věkové struktury mužů a žen Jihočeského kraje zpracovaná z dlouhé výchozí časové řady k 1. 1. se spolehlivostí 95 %, roky 2017, 2026 a 2036.....	112
---------	---	-----

## SEZNAM OBRÁZKŮ

Obr. 1	Schematické vyjádření transformace skutečnosti v model .....	18
Obr. 2	Postup zpracování populačních prognóz .....	19
Obr. 3	Zobrazení kohortně komponentní metody pro populaci žen v Lexisově diagramu.....	24
Obr. 4	Znázornění longitudinálních úmrtnostních tabulek v Lexisově diagramu pro populaci žen.....	26
Obr. 5	Znázornění středního stavu žen v Lexisově diagramu pro výpočet narozených .....	29
Obr. 6	Znázornění migračního salda žen v nejnižších věcích v Lexisově diagramu .....	33
Obr. 7	Znázornění migračního salda žen v nejvyšších věcích v Lexisově diagramu .....	36
Obr. 8	Úprava metody singulárního rozkladu při snížení dimenze diagonální matice $\Sigma$ .....	41
Obr. 9	Definice stochastického procesu na fiktivních hodnotách časové řady.....	55
Obr. 10	Simulace lineárního trendu na fiktivních hodnotách časové řady .....	60
Obr. 11	Odstranění trendu prostřednictvím první difference a následné centrování nově vzniklých hodnot časové řady na nulový průměr, tzv., stacionarizace časové řady.....	61
Obr. 12	Vývoj faktorového skóre třetího faktoru úmrtnosti mužů mezi lety 2002–2014 a jeho prognóza do roku 2024, prognózováno prostřednictvím logaritmického trendu se spolehlivostí 95 % .....	69
Obr. 13	Vývoj faktorového skóre druhého faktoru úmrtnosti žen mezi lety 2002–2014 a jeho prognóza do roku 2024, prognózováno prostřednictvím logaritmického trendu se spolehlivostí 95 % .....	70
Obr. 14	Očekávaný věkový profil pravděpodobností úmrtí žen v roce 2024 zpracovaný modelem Heligmana-Pollarda se spolehlivostí 95 % .....	70
Obr. 15	Očekávaný věkový profil pravděpodobností úmrtí mužů v roce 2024 zpracovaný modelem Heligmana-Pollarda se spolehlivostí 95 % .....	71
Obr. 16	Vývoj naděje dožití chlapců a dívek při narození mezi lety 2002–2014 a jejich prognóza zpracovaná z krátké výchozí časové řady do roku 2024 se spolehlivostí 95 %.....	71
Obr. 17	Vývoj faktorového skóre prvního faktoru u plodnosti mezi lety 2000–2015 a jeho prognóza do roku 2024, prognózováno prostřednictvím logaritmického trendu se spolehlivostí 95 % .....	73
Obr. 18	Očekávané míry plodnosti podle věku v roce 2024 zpracované prostřednictvím Beta funkce se spolehlivostí 95 % .....	74

Obr. 19	Vývoj úhrnné plodnosti mezi lety 2000–2015 a její prognóza zpracovaná z krátké výchozí časové řady do roku 2024 se spolehlivostí 95 % .....	74
Obr. 20	Porovnání skutečného a odhadnutého celkového migračního salda mezi lety 2003–2014 .....	75
Obr. 21	Odhad migračního salda mužů podle jednotek věku, prognózováno prostřednictvím dolního kvartilu, mediánu a horního kvartilu mezi lety 2015–2024.....	76
Obr. 22	Odhad migračního salda žen podle jednotek věku, prognózováno prostřednictvím dolního kvartilu, mediánu a horního kvartilu mezi lety 2015–2024.....	77
Obr. 23	Porovnání prognózy věkové struktury mužů a žen k 1. 1. 2020 oproti 1. 1. 2015 zpracované z krátké výchozí časové řady.....	78
Obr. 24	Porovnání prognózy věkové struktury mužů a žen k 1. 1. 2025 oproti 1. 1. 2015 zpracované z krátké výchozí časové řady.....	78
Obr. 25	Vývoj celkového počtu mužů k 1. 1. mezi lety 2003–2015 a jeho prognóza zpracovaná z krátké výchozí časové řady do roku 2025.....	79
Obr. 26	Vývoj celkového počtu žen k 1. 1. mezi lety 2003–2015 a jeho prognóza zpracovaná z krátké výchozí časové řady do roku 2025 .....	80
Obr. 27	Akaikeho informační kritérium při volbě vyrovnávací funkce u mužů mezi lety 1986–2015 .....	83
Obr. 28	Akaikeho informační kritérium při volbě vyrovnávací funkce u žen mezi lety 1986–2015 .....	83
Obr. 29	Vývoj indexu úrovně úmrtnosti mužů mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno prostřednictvím ARIMA(0,1,1) s driftem.....	84
Obr. 30	Vývoj indexu úrovně úmrtnosti žen mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno prostřednictvím ARIMA(2,1,0) s driftem.....	85
Obr. 31	Očekávaný věkový profil pravděpodobností úmrtí mužů a žen v roce 2016 se spolehlivostí 95 % .....	85
Obr. 32	Očekávaný věkový profil pravděpodobností úmrtí mužů a žen v roce 2035 se spolehlivostí 95 % .....	86
Obr. 33	Vývoj naděje dožití chlapců a dívek při narození mezi lety 1986–2015 a jejich prognóza do roku 2035 se spolehlivostí 95 % .....	87
Obr. 34	Vývoj naděje dožití mužů a žen v 65 letech mezi lety 1986–2015 a jejich prognóza do roku 2035 se spolehlivostí 95 %.....	88
Obr. 35	Očekávané příspěvky ke změně naděje dožití při narození podle jednotek věku mezi lety 2015–2035 .....	89
Obr. 36	Vývoj indexu úrovně plodnosti mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno exponenciální a zpětnou logaritmickou transformací prostřednictvím ARIMA(0,2,1).....	90
Obr. 37	Očekávané míry plodnosti podle věku v roce 2016 se spolehlivostí 95 %.....	91
Obr. 38	Očekávané míry plodnosti podle věku v roce 2035 se spolehlivostí 95 %.....	92

Obr. 39	Vývoj úhrnné plodnosti mezi lety 1986–2015 a její prognóza do roku 2035 se spolehlivostí 95 % .....	92
Obr. 40	Vývoj počtu narozených mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	93
Obr. 41	Vývoj indexu úrovně imigrace mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem .....	94
Obr. 42	Vývoj indexu úrovně imigrace žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem .....	95
Obr. 43	Očekávaný počet přistěhovalých mužů podle věku v roce 2035 se spolehlivostí 95 % .....	96
Obr. 44	Očekávaný počet přistěhovalých žen podle věku v roce 2035 se spolehlivostí 95 % ..	96
Obr. 45	Vývoj počtu přistěhovaných mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	97
Obr. 46	Vývoj počtu přistěhovaných žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	98
Obr. 47	Vývoj indexu úrovně emigrace mužů mezi lety 1994–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem .....	99
Obr. 48	Vývoj indexu úrovně emigrace žen mezi lety 1994–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem .....	99
Obr. 49	Očekávané míry vystěhování mužů podle věku v roce 2035 se spolehlivostí 95 % ..	100
Obr. 50	Očekávané míry vystěhování žen podle věku v roce 2035 se spolehlivostí 95 % .....	100
Obr. 51	Vývoj počtu vystěhovalých mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	101
Obr. 52	Vývoj počtu vystěhovalých žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	101
Obr. 53	Vývoj migračního salda mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	102
Obr. 54	Vývoj migračního salda žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	103
Obr. 55	Porovnání prognózy věkové struktury mužů a žen k 31. 12. 2016 oproti 31. 12. 2015 s 95% spolehlivostí .....	104
Obr. 56	Porovnání prognózy věkové struktury mužů a žen k 31. 12. 2025 oproti 31. 12. 2015 s 95% spolehlivostí .....	104
Obr. 57	Porovnání prognózy věkové struktury mužů a žen k 31. 12. 2035 oproti 31. 12. 2015 s 95% spolehlivostí .....	105
Obr. 58	Vývoj celkového počtu mužů k 1. 1. mezi lety 1986–2016 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	106
Obr. 59	Vývoj celkového počtu žen k 1. 1. mezi lety 1986–2016 a jeho prognóza do roku 2035 se spolehlivostí 95 % .....	106

## SEZNAM POUŽITÝCH ZKRATEK

PřF UK	Přírodovědecká fakulta Univerzity Karlovy
VÚV TGM	Výzkumný ústav vodohospodářský T. G. Masaryka
ČSÚ	Český statistický úřad
SVD	Singulární rozklad matic
Níz.	Nízká varianta prognózy
Stř.	Střední varianta prognózy
Vys.	Vysoká varianta prognózy

## Kapitola 1

### Úvod

Demografie zachycuje radostné a smutné okamžiky v životě člověka. Pomáhá zlepšit životní podmínky a podchytit závažné populační a celospolečenské problémy. Demografie se zabývá minulostí, zachycuje současné trendy a vyjadřuje očekávání o budoucnosti. Právě budoucnost se stala základem této práce.

*„(...) Populace, když se nezmění, roste v geometrickém poměru. Obživa roste pouze v aritmetickém poměru. (...)“ (Malthus, 1798, s. 4).*

Tímto výrokem vyjádřil Thomas Robert Malthus obavu před přelidněním populace na konci 18. století. Matematicky vyjádřil očekávání o budoucím směřování velikosti populace a množství obživy a jejich možných dopadech. Stejným způsobem jako se Malthus snažil nastítnit matematicky budoucí vývoj populace, bude i v této práci kladen důraz na analytické zpracování očekávaného vývoje populace s přihlédnutím k současným trendům v oblasti populačního prognózování. Tato diplomová práce se snaží o sumarizaci současných nejpoužívanějších metod v oblasti populačních prognóz a jejich aplikaci na menší územní celek Jihočeského kraje.

V současnosti zpracovávají populační prognózy v České republice tři instituce: Český statistický úřad, Katedra demografie a geodemografie Přírodovědecké fakulty Univerzity Karlovy a Katedra demografie Fakulty informatiky a statistiky Vysoké školy ekonomické. Český statistický úřad uveřejnil v roce 2004 populační prognózu České republiky do roku 2050 (ČSÚ, 2004d) a v roce 2013 zpracoval projekci České republiky do roku 2100 (ČSÚ, 2013e). V roce 2014 vytvořil Český statistický úřad projekci obyvatel v krajích České republiky včetně Jihočeského kraje do roku 2050 (ČSÚ, 2014e). Na Katedře demografie a geodemografie Přírodovědecké fakulty zpracovávají populační prognózy RNDr. Tomáš Kučera, CSc. a RNDr. Boris Burcin, Ph.D. Vypracovali prognózu České republiky od roku 2003 do roku 2065 (Burcin a Kučera, 2004, s. 100–111). V roce 2011 vytvořili populační prognózu České republiky do roku 2070 (Burcin a Kučera, 2011, s. 181–212). Na Katedře demografie Fakulty informatiky a statistiky Vysoké školy ekonomické bývají prognózy zpravidla spojeny s odvozenými prognózami, na kterých se podílí doc. Ing. Jitka Langhamrová, CSc. a RNDr. Tomáš Fiala, CSc. Příkladem je prognóza lidského kapitálu obyvatelstva České republiky do roku 2050

(Langhamrová a kol., 2010, s. 181–196). Společně zpracovali projekci obyvatel v krajích ČR (Langhamrová a Fiala., 2014, s. 73–96).

Projekci Jihočeského kraje vytvořila ve své diplomové práci Bartková (2009), kterou zpracovala v programu Cohort Component Projections. Očekávaný vývoj počtu obyvatel prostřednictvím trendových funkcí zpracovala ve své bakalářské práci Cahová (2012, s. 31). Pro Jihočeský kraj byla vytvořena řada odvozených prognóz, kde příkladem je vývoj kapacit v domovech pro seniory (Šídllová Kunstová a Šídllo, 2016, s. 113–122). Vzhledem ke skutečnosti, že byla zpracována velice kvalitní prvotní analýza Jihočeského kraje (2013), nebude podrobněji v této práci popsána a více bude kladen důraz na popis metod a výsledků populační prognózy Jihočeského kraje.

## 1.1 Cíle práce

Tato diplomová práce vznikla za účelem představení a aplikace současných trendů a vybraných modelů v tvorbě populačních prognóz. Vzhledem k osobnímu vztahu autora byl zvolen pro aplikaci populační prognózy Jihočeský kraj. Výběr konkrétního modelu pro tvorbu populační prognózy poté závisí na mnoha okolnostech. Jedním z limitujících faktorů je délka časové řady vstupních dat, která je k dispozici. Josef Arlt a Markéta Arltová (2003, s. 9) tento problém nazývají již v předmluvě datová nedostatečnost, kdy na základě odlišné délky časové řady je nutné volit rozdílné modelovací postupy. Právě tato myšlenka dala podnět pro vznik celé diplomové práce. Pro ilustraci odlišných přístupů bude zvolena krátká výchozí časová řada a dlouhá výchozí časová řada, a tím bude ukázán jejich vliv na prognózu.

Prvním cílem je popis vybraných metod pro tvorbu populačních prognóz. Pro naplnění tohoto cíle jsou zvoleny klíčové publikace autorů Shryocka a kol. (1973), Pavlíka a kol. (1986), Klostermana (1990), Prestona a kol. (2001), dále článek McCAA (1989) a disertační práce Kučery (1998). Parciálním cílem je charakteristika vybraných demografických modelů pro tvorbu dílčích prognóz úmrtnosti, plodnosti a migrace. Aby byl naplněn tento cíl, byla vybrána základní literatura Pressata (1995), Hindeho (1998) a články Heligmana a Pollarda (1980), Leea a Cartera (1992), Leea (1993) a Raymera a kol. (2012).

Druhým cílem je vytvoření adekvátní prognózy Jihočeského kraje z krátké výchozí časové řady. Smyslem je simulovat méně kvalitní a špatně dostupné údaje pro malé územní celky nebo pro rozvojové země na datech Jihočeského kraje a navrhnout možné řešení zpracování dílčích prognóz prostřednictvím analytických funkcí. Z toho vyplývá několik parciálních cílů, mezi které patří výběr vhodného demografického modelu a funkce pro dílčí prognózu úmrtnosti, plodnosti a migrace z této krátké výchozí časové řady. Vzhledem ke skutečnosti, že data za migraci nejsou často k dispozici, je záměrem tato data odhadnout.

Třetím cílem je zpracování kvalitní populační prognózy Jihočeského kraje s 95% a 75% spolehlivostí z dlouhé výchozí časové řady. Parciálním cílem je vytvoření dílčích prognóz úmrtnosti, plodnosti a migrace prostřednictvím Lee-Carterových modelů analýzou dlouhé výchozí časové řady. Mezi individuální cíle poté patří volba vhodné extrapolační metody při

analýze odhadnutého indexu úrovně úmrtnosti, plodnosti, emigrace a imigrace z příslušného Lee-Carterova modelu.

Posledním čtvrtým cílem je vzájemné porovnání populační prognózy Jihočeského kraje vytvořené z dlouhé výchozí časové řady s populační projekcí Jihočeského kraje zpracovanou Českým statistickým úřadem (ČSÚ, 2014e) a střední variantou prognózy Jihočeského kraje vytvořenou v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka (Burcin a kol., 2014). Parciálním cílem je vzájemné porovnání dílčích prognóz úmrtnosti, plodnosti a migrace vytvořených z dlouhé výchozí časové řady s jejich odbornými odhady zpracovanými oběma zmíněnými institucemi.

## 1.2 Struktura práce

Tato diplomová práce se skládá z devíti kapitol. V první kapitole jsou vytyčeny základní cíle práce. Druhá kapitola se věnuje vymezení populačních prognóz. Jsou popsány základní pojmy charakterizující populační prognózy, které se objeví později v textu. Dále se stanoví fáze zpracování populačních prognóz. Třetí kapitola popisuje vybrané metody tvorby populačních prognóz. Jsou charakterizovány vybrané analytické modely a funkce, které se v dnešní době nejčastěji používají při zpracování dílčích prognóz.

Čtvrtá kapitola se věnuje analýze časových řad. Jsou charakterizovány vybrané trendové funkce a hlouběji je popsána Box-Jenkinsonova metodologie. Pátá kapitola se zaměřuje na zpracování datových zdrojů a představuje rozdělení těchto dat na krátkou výchozí časovou řadu a dlouhou výchozí časovou řadu. V šesté kapitole je představeno jedno z možných řešení zpracování prognózy Jihočeského kraje z krátké výchozí časové řady do 1. 1. 2025. Je poukázáno na problémy, které s analýzou krátké časové řady souvisí. V sedmé kapitole je vytvořena plnohodnotná prognóza Jihočeského kraje prostřednictvím Lee-Carterových modelů z dlouhé výchozí časové řady do 31. 12. 2035. Osmá kapitola porovnává prognózu Jihočeského kraje z dlouhé výchozí časové řady se střední variantou prognózy Jihočeského kraje zpracovanou v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka (Burcin a kol., 2014) a s projekcí Českého statistického úřadu (ČSÚ, 2014e). Devátá a závěrečná kapitola shrnuje výsledky této diplomové práce.



## Kapitola 2

### Vymezení populačních prognóz

Vytváření prognóz patří k výsledkům demografického studia. Pro jejich tvorbu je potřebný značný teoretický základ a mnoho zkušeností. Tato kapitola se soustředí na teoretické zázemí a vymezení fází přípravy populačních prognóz, kdy nejprve budou charakterizovány základní pojmy a poté vymezeno samotné prognózování.

#### 2.1 Základní pojmy charakterizující populační prognózy

Na počátku je nezbytné vymezit rozdíl mezi pojmy prognóza a projekce. Často jsou obě slova užívána nesprávně jako synonymum. Nejprve k prvnímu pojmu **prognóza**. Demografové definují prognózu jako „(...) *nepodmíněnou, na vědeckém poznání založenou výpověď o očekávaném a v době jejího vzniku nejpravděpodobnějším budoucím vývoji sledovaného jevu* (...)“ (Kučera, 1998, s. 17). Statisticky by bylo možné definovat prognózu jako nejpravděpodobnější množinu nejistých hodnot. **Projekce** „(...) *představuje výpověď o perspektivním vývoji, který je důsledkem naplnění libovolných předpokladů bez explicitního nároku na jejich reálnost* (...)“ (Kučera, 1998, s. 17).

Dále je potřeba vymezit význam pojmů práh projekce, horizont projekce a krok projekce. **Práh projekce** představuje okamžik, od kterého bude projekce sestrojena. **Horizont projekce** ztělesňuje kýžený bod, ke kterému pak směřuje celé úsilí. Je zřejmé, že čím více je horizont vzdálen od jeho prahu, tím se zvyšuje nejistota předpovědi. Tato vzdálenost se označuje jako *délka projekce*. **Krok projekce** rozděluje délku projekce do pravidelných intervalů. Nejčastěji je zvolen krok projekce jako jednoletý, pětiletý případně desetiletý. Velikost kroku projekce závisí také na výchozí věkové struktuře. Pokud jsou k dispozici jednoleté věkové skupiny lze je rozdělit například na pětiletý krok projekce. Pro pětileté věkové skupiny ale nelze sestrojit jednoletý krok projekce (Štědroň a kol., 2012, s. 96–97).

Prognózy lze podle jejich délky dělit na krátkodobé, střednědobé a dlouhodobé. **Krátkodobé prognózy** „(...) *jsou vypočítány zpravidla na období kratší než jedna generace*, (...)“ **střednědobé prognózy** jsou vymezeny „(...) *na období přibližně jedné generace* (...)“ a **dlouhodobé prognózy** se zpracovávají „(...) *na delší časový úsek* (...)“ (Pavlík a Kalibová, 2005, s. 101). Kučera (1998, s. 20) dělí prognózy také z hlediska územního rozsahu na národní

a regionální. Tento způsob lze rozšířit o globální a místní prognózy. **Globální prognózy** se týkají celého světa, případně jeho části. **Národní prognózy** se omezují na populaci státu. **Regionální prognózy** zohledňují členění státu. Vytváří se například krajské nebo okresní prognózy. Poslední skupinou prognóz jsou **místní prognózy**. Jedná se například o prognózy obcí.

Burcin a kol. (2007, s. 12) dělí prognózy na základní, elementární a odvozené. **Základní prognózy** jsou charakterizovány strukturou a početními stavy obyvatelstva. **Elementární prognózy** tvoří podklad pro základní prognózy a jsou reprezentovány dílčími prognózami úmrtnosti, plodnosti a migrace. Základní a elementární prognózy tvoří dohromady velkou skupinu populačních prognóz. **Odvozené prognózy** se zpracovávají ze základní prognózy. Příkladem mohou být prognózy struktury domácností nebo prognózy počtu žáků.

**Obr. 1:** Schematické vyjádření transformace skutečnosti v model



**Zdroj:** Fábry (2010, s. 7).

Dalším důležitým pojmem při tvorbě prognóz je **model**. V matematickém modelování představuje model zjednodušení reality, viz obrázek 1. Problém spočívá ve výběru správného modelu. Příliš zjednodušený model bude vykazovat nereálné výsledky. Naopak velmi složitý model bude neproveditelný a výsledky nebudou dosažitelné v reálném čase (Fábry, 2010, s. 7). V demografii se modelem označují převážně tabulkové modely, které určitým způsobem vyjadřují řád populačního vývoje (Pavlík a Kalibová, 2005, s. 101–102). Příkladem tabulkových modelů mohou být úmrtnostní tabulky, tabulky sňatečnosti, tabulky rozvodovosti apod.

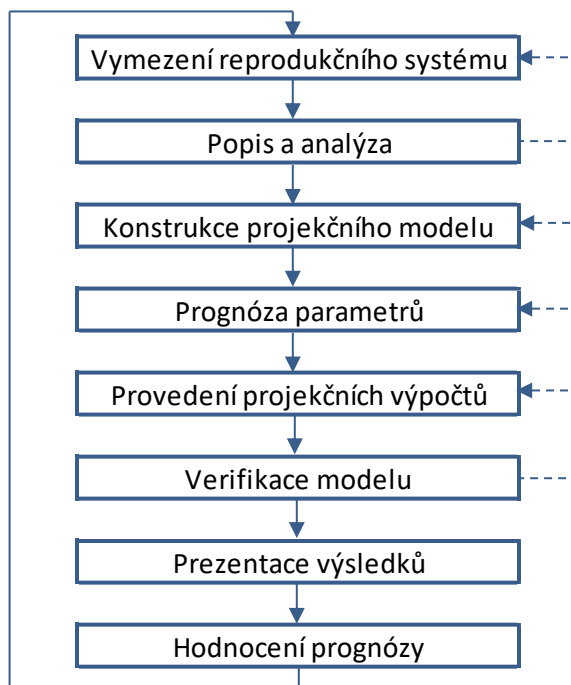
Rozlišují se deterministické a stochastické modely. **Stochastické modely** jsou charakterizovány pravděpodobností „(...) nastoupení určitých událostí u jednotlivců během studovaného období (...)“ a **deterministické modely** „(...) jsou založeny na funkčních vztazích mezi jednotlivými proměnnými a předpokládají nekonečnou velikost populací (...)“ (Pavlík a Kalibová, 2005, s. 101).

## 2.2 Obecná formulace prognózování a stanovení fází zpracování populačních prognóz

Prognózování představuje souhrnný proces pro formování prognóz. Jedná se o souhrn činností, jejichž cílem je podání informace o budoucím vývoji zkoumaného subjektu. Jednotlivé činnosti na sebe navazují a utvářejí celý systém (Kučera, 1998, s. 18). Prognózování se využívá v řadě oborů například v ekonomii, managementu, ale také v demografii. Předmětem studia prognózování je v demografii obyvatelstvo. Prognózují se nejen celkové počty, ale především

struktura obyvatelstva, která se často ještě člení podle různých znaků například pohlaví, vzdělání, ekonomické aktivity, národnosti a další (Štědroň a kol., 2012, s. 95).

*Obr. 2: Postup zpracování populačních prognóz*



**Zdroj:** Kučera (1998, s. 19) a Fábry (2010, s. 10).

Prognózování se v demografii skládá z osmi etap viz obrázek 2. První etapou je **vymezení reprodukčního systému**. Detekují se významné prvky a vazby v systému. Cílem druhé etapy je **popis a analýza** reprodukčního systému (Kučera, 1998, s. 18). Mezi první a druhou etapou existuje silná vazba. Pokud je vymezení reprodukčního systému nedostatečné a nelze provést analýzu, je nutné zpětně dodefinovat reprodukční systém. Obecně je nutné vymezit a popsat sledované území, definovat práh, horizont a krok prognózy, určit výchozí okamžik, ke kterému bude prognóza stanovena – zpravidla k 1. 1., 1. 7. nebo 31. 12., analyzovat pohlavně věkovou strukturu a další.

Třetí etapou je **konstrukce projekčního modelu**. Definuje se metoda, která bude užita k budoucí předpovědi. Jedná se například o metody extrapolace nebo kohortně komponentní metody. Další etapou je **prognóza parametrů** (Kučera, 1998, s. 18–19). Kalkulují se elementární prognózy, kdy se odhaduje vývoj intenzit úmrtnosti, plodnosti a migrace. Štědroň a kol. (2012, s. 3) doporučují důsledný rozbor minulých pozorování těchto intenzit, které by se měly promítnout do očekávaného vývoje populace. Zároveň je ale také nutné sledovat aktuální populační klima a jeho budoucí směřování, více viz Kocourková (2007, s. 95–102). Pátá etapa spočívá v **provedení projekčních výpočtů** prostřednictvím počítačových programů a je ukončena dokumentací výsledků (Kučera, 1998, s. 19). Fábry (2010, s. 12) v této souvislosti zařazuje další důležitý prvek, kterým je **verifikace modelu**. Tato etapa by měla být přítomná od konstrukce projekčního modelu po provedení projekčních výpočtů. Jedná se o posouzení vhodnosti zvoleného modelu a reálnosti dosažených výsledků. Je potřeba si klást otázku, zda by

jiný model neposkytoval kvalitnější výsledky, případně jestli by jiný software neurychlil zpracování prognózy.

**Prezentace výsledků** patří k sedmé etapě postupu zpracování populačních prognóz. Zvolí se podstatné informace a představí se uživateli. Poslední osmou etapou je **hodnocení prognózy**. Cílem této etapy je porovnání výsledků se skutečností. Výsledky hodnocení by se měly promítnout při aktualizaci staré nebo při tvorbě nové prognózy (Kučera, 1998, s. 19).

## Kapitola 3

### Vybrané metody tvorby populačních prognóz včetně vybraných modelů směřujících k vytvoření dílčí prognózy úmrtnosti, plodnosti a migrace

Tato kapitola se zabývá popisem konstrukce populačních prognóz. Existuje celá řada metod, které se věnují výpočtu populačních prognóz. K nejjednodušším patří metody **extrapolace** celkového počtu obyvatel. Pro odhad budoucího vývoje se používají trendové funkce, které s určitou spolehlivostí vyjadřují budoucí populační vývoj. Tyto metody ale neříkají nic o struktuře obyvatelstva. Z tohoto důvodu byly extrapolací metody rozšířeny o odhad věkové struktury (Pavlík a kol., 1986, s. 389). Další velkou skupinou metod pro výpočet populačních prognóz jsou **kohortně komponentní metody**. Tyto metody jsou navrženy pro výpočet populačních projekcí, ale při zařazení mezi spolehlivosti dílčích prognóz do výpočtu, je jejich potenciál rozšířen i pro populační prognózy. Kohortně komponentní metody patří mezi nejčastěji aplikované metody při odhadu budoucích struktur (Štědroň a kol., 2012, s. 97; Kučera, 1998, s. 24–25). Rozebrána zde bude kohortně komponentní metoda bez migrace a kohortně komponentní metoda s migrací.

Aby bylo možné sestavit populační prognózu, je nutné nejprve vytvořit dílčí prognózy úmrtnosti, plodnosti a migrace. Tyto prognózy lze stanovit prostřednictvím odborného odhadu, který vyžaduje značnou zkušenost a detailní popis očekávaného vývoje demografických a ekonomických podmínek. Další možností je odhadnout budoucí vývoj úmrtnosti, plodnosti a migrace na základě minulého průběhu časové řady (Pavlík a kol., 1986, s. 396). Tímto způsobem budou vytvořeny dílčí prognózy v této práci.

Existuje velké množství demografických modelů, zákonitostí a funkcí, jejichž cílem je vyjádřit rozložení intenzit úmrtnosti, plodnosti a migrace podle věku, popsat je co nejpřesněji a s co nejmenším počtem parametrů. Tyto parametry lze poté extrapolovat do budoucnosti například prostřednictvím sledování trendů jejich minulého vývoje (Gavrilová a Gavrilov, 2011, s. 110). Zde více přiblížené modely, zákonitosti a funkce byly zvoleny, protože vystihují rozložení intenzit úmrtnosti, plodnosti a migrace v každém věku.

Pro dílčí prognózu úmrtnosti byly zvoleny dva modely, jmenovitě Heligman-Pollardův model úmrtnosti a Lee-Carterův model úmrtnosti. Tyto modely jsou popsány v podkapitole 3.2. **Heligman-Pollardův model úmrtnosti** se častěji aplikuje v zahraničních pracích, příkladem je studie Keilmana a kol. (2001, s. 50–58). Jedná se o jednofaktorový model, který v sobě zahrnuje

pouze faktor věku. Model převážně slouží k vyrovnání křivky kvocientů úmrtnosti, ale časová řada jeho odhadnutých parametrů může být extrapolována do budoucnosti. Druhým modelem užitým v této souvislosti bude **Lee-Carterův model úmrtnosti**, který se řadí mezi dvoufaktorové modely. Jednak zahrnuje faktor věku a také obsahuje faktor času (Štádlarová, 2013, s. 14–15). Lee-Carterův model úmrtnosti se tradičně používá jak v zahraničním, tak i v českém prostředí, kde příkladem je článek Habartové (2009).

Pro vytvoření dílčí prognózy plodnosti jsou vybrány dva postupy zpracování intenzit plodnosti podle věku. První z nich je stanoven prostřednictvím Beta funkce a druhý pomocí Lee-Carterova modelu plodnosti. Oba postupy jsou zpracovány v podkapitole 3.3. Analýza prostřednictvím **Beta funkce** patří spíše k netradičním způsobům tvorby prognóz. Častěji se užívá Gama funkce, kterou popisují Keilman a kol. (2001, s. 23–35). Beta funkce je klasifikována jako jednofaktorový model plodnosti, kde faktorem je věk. Slouží převážně pro vyrovnání křivek plodnosti, ale její parametry mohou být rovněž extrapolovány do budoucnosti. Druhým modelem užívaným při odhadu rozložení intenzit plodnosti podle věku je **Lee-Carterův model plodnosti**. Jedná se o dvoufaktorový model, který má v sobě zahrnut faktor věku a faktor času. Lee-Carterův model plodnosti použila v českém prostředí například Pohořálková (2007, s. 26–29). Oba modely budou v této kapitole vysvětleny.

Jako poslední bude v podkapitole 3.4 popsána metodika zpracování dílčí prognózy migrace. Protože často nejsou k dispozici věkově specifické údaje o migraci, bude představena **nepřímá metoda výpočtu migračního salda**. Při konstrukci předpovědi bude poté uvažováno konstantní migrační saldo, které bude prognózováno 25%, 50% a 75% kvantilem odhadnutého migračního salda zvlášť podle jednotek věku. V této souvislosti lze také užít Lee-Carterův model čisté migrace (García-Guerrero, 2016, s. 119). Pokud jsou k dispozici podrobnější data za migraci, lze ji poté analyzovat zvlášť za přistěhovalé **Lee-Carterovým modelem imigrace** a zvlášť za vystěhovalé **Lee-Carterovým modelem emigrace** (Raymer a kol., 2012, s. 1).

### 3.1 Kohortně komponentní metoda

Význam kohortně komponentní metody spočívá v zahrnutí komponent (úmrtnosti, plodnosti a případně i migrace) do projekce obyvatel (Swanson a Tayman, 2012, s. 200). Rozlišují se dva typy této metody, a to kohortně komponentní metoda bez migrace, kde komponentami jsou úmrtnost a plodnost. Druhým typem je kohortně komponentní metoda s migrací, která je rozšířena o komponentu migrace. Toto rozdělení používá Pavlík a kol. (1986, s. 389), ale i King a Connell (1999, s. 87). Pavlík a Kalibová (2005, s. 19) definují pojem kohorta jako soubor osob, kteří prožili v daný okamžik určitou událost.

Prostředky k odhadu populačního vývoje jsou známy již od počátku 19. století. Definovány byly již míry plodnosti a úmrtnosti podle věku, ale nebyly využity pro potřeby prognóz téměř do počátku 20. století (Smith a kol., 2013, s. 193). Jako první se kohortně komponentní metodou zabýval Edwin Cannan (1895). V minulosti se růst populace odhadoval převážně extrapolací křivek stavů obyvatel, případně expertním odhadem těchto stavů. Populační vývoj byl vnímán jako celek bez vnitřní diferenciací, kdy nedocházelo ke spojení jednotlivých komponent

s vývojem populačním, tzn. úmrtnost, plodnost a migrace byly studovány zvlášť, nikoli jako jednotlivé složky populačního vývoje (Kučera, 1998, s. 21). Cannan nastínil názor o přínosu zařazení komponent do populačního vývoje a tvrdí, že „(...) skutečnou otázkou není, zda bychom se měli úplně zdržet od odhadování růstu populace, ale zda bychom se měli spokojit s odhady, které byly vytvořeny bez adekvátního zvážení všech dat, které jsou k dispozici (...)“ (Cannan, 1895, s. 198). Kohortně komponentní metodu aplikoval také Bowley (1924), který ji zpracoval nezávisle na práci Cannana (Smith a kol., 2013, s. 193). Dále se touto metodou zabýval Whelpton (1928), který navázal na práci svého předchůdce Bowleyho (Kučera, 1998, s. 21).

Kohortně komponentní metoda se používá pro odhad budoucí věkové struktury a počítá se odděleně podle pohlaví (Pavlík a kol., 1986, s. 385). Princip metody spočívá v posunu osob v rámci jedné generace do vyššího věku za stálého působení úmrtnosti, plodnosti a migrace (Štědroň a kol., 2012, s. 97). Pro další analýzu bude uvažován jednoletý krok projekce. Vyšším krokem projekce se zabývá například Aubrecht (2016, s. 19–20) nebo Štědroň a kol. (2012, s. 98). Hranice posledního otevřeného intervalu mohou být různé, pro ilustraci postupu je v této práci uveden poslední otevřený interval ve věku 100 a více let.

Pavlík a kol. (1986, s. 397) doporučují jako úplně první krok při sestavování populačních projekcí stanovit a kvantifikovat hypotézu budoucího vývoje úmrtnosti, plodnosti a migrace. Po kvantifikaci těchto jednotlivých hypotéz následuje aplikace jedné z metod populačních projekcí například kohortně komponentní metoda.

### 3.1.1. Kohortně komponentní metoda bez migrace

Do odhadu budoucího vývoje je v rámci kohortně komponentní metody bez migrace jako komponenta zahrnuta úmrtnost, plodnost, ale abstrahuje se od migrace. Jedná se o tzv. uzavřenou populaci. Kohortně komponentní metoda je založena na bilanční rovnici. Pro nejjednodušší případ, kdy se do výpočtu nezahrnuje migrace, vypadá rovnice (Weinstein a Pillai, 2016, s. 256)

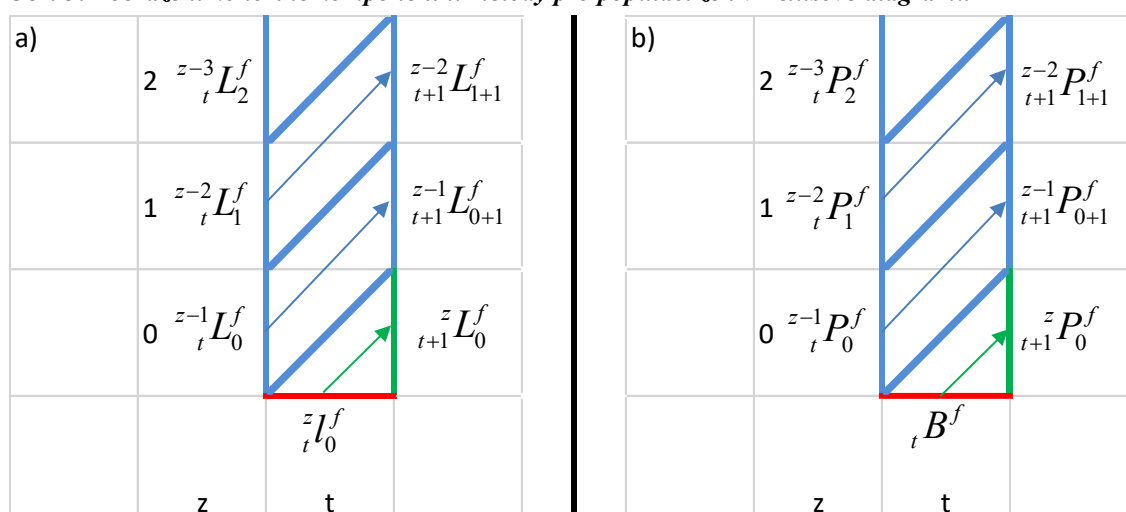
$$P_t = P_0 + B - D. \quad (3.1)$$

Na levé straně populační rovnice (3.1) je celkový počet obyvatel na konci roku  $t$  pod zkratkou  $P_t$ , který se vypočítá z výchozího stavu celkového počtu obyvatel označeného  $P_0$ , ke kterému se přičtou narození od výchozího okamžiku do konce roku  $t$  pod značkou  $B$  a odečtou se zemřelí rovněž od výchozího okamžiku do konce roku  $t$  pod označením  $D$ . Výchozí okamžik je znázorněn jako rok nula.

Aby bylo možné sestavit populační projekci kohortně komponentní metodou bez migrace, je nezbytné mít k dispozici následující data (Pavlík a kol., 1986, s. 397):

- Stavy mužů a žen podle věku k výchozímu okamžiku.
- Počty zemřelých mužů a žen podle věku vyjádřené úmrtnostními tabulkami.
- Míry plodnosti podle věku matky.

Bez těchto dat není možné projekci kohortně komponentní metodou sestavit.

**Obr. 3: Zobrazení kohortně komponentní metody pro populaci žen v Lexisově diagramu**

**Zdroj:** vlastní zpracování na základě schématu Pavlíka a kol. (1986, s. 398).

**Pozn.:** symboly uvedené na obrázku jsou popsány v tabulce 1.

Na obrázku 3 v kategorii *a* je naznačen princip komponentní metody v úmrtnostní tabulce žen. Jsou zobrazeny tabulkové počty žijících žen v dokončeném věku  $x$  k 1. 1. roku  $t$  z generace  $z - l - x$ , kde index pro kohortu se zobrazuje u vzorce vlevo nahoře. Graficky je modrou barvou znázorněn posun těchto žijících na počátek dalšího roku, tedy k 1. 1. roku  $t + l$ , kdy generace zůstává samozřejmě stejná a věk se o rok zvýší na  $x + l$ . Generací  $z$  jsou označeny tabulkové počty dožívajících v přesném věku nula v roce  $t$ , které se v obrázku znázorňují červenou barvou. Následně dochází k jejich posunu do tabulkového počtu žijících žen v dokončeném věku nula k 1. 1. roku  $t + l$  z generace  $z$ . Schematicky je tento posun zobrazen zelenou barvou. V kategorii *b* je již ilustrován výsledek kohortně komponentní metody na skutečné věkové struktuře. Modrou barvou jsou znázorněny skutečné počty žen ve věku  $x$  k 1. 1. roku  $t$  z generace  $z - l - x$  a při aplikaci pravděpodobností přežití žen v dokončeném věku u uzavřené populace dochází k jejich posunu o rok déle, tedy  $t + l$ , kdy generace zůstávají přirozeně stejné a věk se zvýší na  $x + l$ . Živě narozené dívky v roce  $t$  se přesunou do výchozího okamžiku 1. 1. roku  $t + l$  generace  $z$  při působení kojenecké úmrtnosti u uzavřené populace. Tento moment je schematicky zobrazen zelenou barvou.

Z obrázku 3 kategorie *a* i *b* vychází princip kohortně komponentní metody, kdy je naznačen posun osob v rámci jedné generace  $z$  1. 1. roku  $t$  k 1. 1. roku  $t + l$ . Obrázek 3 slouží také jako základ grafického zobrazení výpočtu kohortně komponentní metody bez migrace. Zpracování vzorců je navíc rozšířeno o generace, ve kterých jsou dobře viditelné rozdíly mezi transversálním a longitudinálním přístupem. Transverzální pohled nabízí značné zjednodušení, které spočívá v analýze věkové struktury k jednomu roku při zahrnutí všech generací. Longitudinální analýza je založena na posouzení jedné generace od narození do věku, kdy nikdo nezůstane naživu (Pavlík a kol., 1986, s. 100). V analytické části této práce bude použit transversální přístup. Pro snadnější orientaci ve výpočtech je v tabulce 1 uveden seznam použitých symbolů. Pokud nebude uvedeno jinak, bude toto značení použito pro celou podkapitolu kohortně komponentní metody bez migrace. Jednotlivé vzorce budou konstruovány tak, že se nejprve sestrojí schéma výpočtu a následně se přiřadí omezující podmínky, pro které bude výraz definován.



Tab. 1: Seznam symbolů použitých u kohortně komponentní metody bez migrace

Symbol	Popis
$f$	ženy
$m$	muži
$x$	dokončený věk
$t$	rok
$z$	generace neboli rok narození $t^*$
$k$	krok transformace
${}^{z-1-x}_t S_x^f$	pravděpodobnost přežití žen v dokončeném věku $x$ od počátečního stavu v roce $t$ ke koncovému stavu v roce $t$ z generace $z - l - x$
${}^{z-1-x}_t L_x^f$	tabulkový počet žijících žen v dokončeném věku $x$ v roce $t$ z generace $z - l - x$
$T_x^f$	celkový počet let zbývajících k dožití tabulkové populaci žen v dokončeném věku $x$ v roce $t$
${}^{z-1-x}_t P_x^{k,f}$	počet žijících žen v dokončeném věku $x$ k výchozímu okamžiku v roce $t$ z generace $z - l - x$ v $k$ -tém kroku transformace
${}^{z-1-x}_t \hat{P}_x^{k,f}$	odhad počtu žijících žen v dokončeném věku $x$ k výchozímu okamžiku v roce $t$ z generace $z - l - x$ v $k$ -tém kroku transformace
${}^{z-1-x}_t B_x^v$	počet živě narozených v roce $t$ ženám z generace $z - l - x$ v dokončeném věku $x$
${}_t \hat{B}^{k+1,f}$	odhad počtu živě narozených dívek v roce $t$ v $k + 1$ kroku transformace
${}_t f_x$	míry plodnosti podle věku $x$ ke střednímu stavu roku $t$ spočítané z třetího hlavního souboru
${}^{z-1-x}_t f$	míry plodnosti z generace žen $z - l - x$ ke střednímu stavu roku $t$ spočítané z druhého hlavního souboru
$\phi$	podíl děvčat při narození
${}_t l_0^f$	kořen tabulek žen v roce $t$ z generace $z$

**Zdroj:** Kučera (1998, s. 75–78), vlastní zpracování.

Obecně lze techniku kohortně komponentní metody bez migrace rozdělit do třech bodů. Prvním z nich je tzv. projekce žijících, která se skládá z výpočtu pravděpodobností přežití v dokončeném věku a jejich aplikace na věkovou strukturu žijících (Roubíček, 1996, s. 219). Druhý bod bývá často spojován s prvním a lze jej popsat jako projekce žijících v nejvyšších věcích. Skládá se z odhadu pravděpodobnosti přežití v posledním otevřeném intervalu a jejich uplatnění na počty žijících v nejvyšších věcích. Poslední bodem je projekce narozených a jejich posun na konec roku. Jednotlivé body budou nyní podrobně rozebrány.

Jak již bylo naznačeno, prvním bodem kohortně komponentní metody bez migrace je tzv. **projekce žijících**. Aby bylo možné odhadnout počty žijících, musí se nejprve vypočítat pravděpodobnosti přežití mužů a žen v dokončeném věku. Tady je možné postupovat dvěma směry, a to buď využít longitudinální nebo transversální přístup. Přesnější výpočet nabízí **longitudinální přístup**, který vychází ze základního vzorce uvedeného v knize Shryocka a kol. (1973, s. 452)

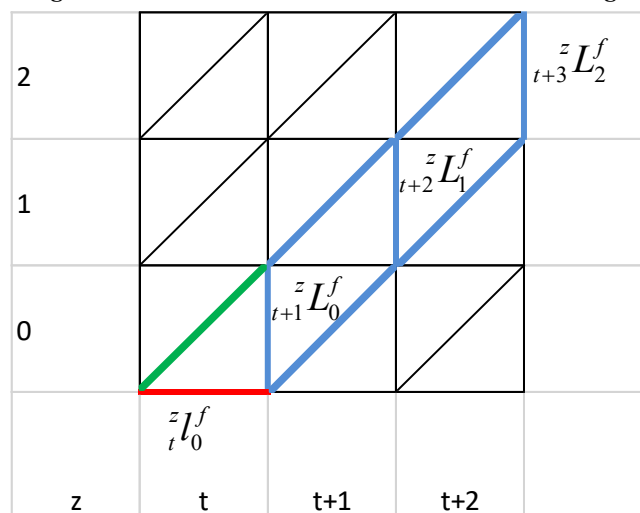
$${}_t S_x^f = \frac{{}_t L_{x+1}^f}{{}_t L_x^f}, \quad x = 0, 1, 2, \dots, 98, \quad (3.2)$$

$${}_t S_x^m = \frac{{}_t L_{x+1}^m}{{}_t L_x^m}, \quad x = 0, 1, 2, \dots, 98. \quad (3.3)$$

Longitudinální přístup je charakteristický tím, že se analyzuje celá jedna kohorta od narození až po úmrtí. Aby bylo možné tento postup aplikovat, jsou zapotřebí longitudinální úmrtnostní

tabulky. Problém těchto tabulek spočívá v tom, že data musí být členěna v druhém hlavním souboru a celá generace musí být vymřelá. Tím je možné analyzovat pouze historická data. Podle vzorce (3.2) se vypočítají pravděpodobnosti přežití žen a na základě formulace (3.3) se vyjádří pravděpodobnosti přežití mužů.

**Obr. 4: Znárodnění longitudinálních úmrtnostních tabulek v Lexisově diagramu pro populaci žen**



**Zdroj:** vlastní zpracování.

Na obrázku 4 je zobrazen princip výpočtu pravděpodobností přežití žen z longitudinálních úmrtnostních tabulek, který spočívá v podílu počtu žijících na konci roku  $t + 1$  a počtu žijících na počátku roku  $t + 1$ . Pro pochopení této problematiky je nutné si uvědomit, že  $zL_{t+2}$  na počátku intervalu roku  $t + 2$  je shodný s  $zL_{t+1}$  na konci intervalu roku  $t + 1$ .

Pravděpodobnosti přežití mužů a žen se při uvažování **transversálního přístupu** vypočítají (Shryock a kol., 1973, s. 452)

$${}^{z-1-x}_t S^f_x = \frac{{}^{z-2-x}_t L^f_{x+1}}{{}^{z-1-x}_t L^f_x}, \quad x = 0, 1, 2, \dots, 98, \quad (3.4)$$

$${}^{z-1-x}_t S^m_x = \frac{{}^{z-2-x}_t L^m_{x+1}}{{}^{z-1-x}_t L^m_x}, \quad x = 0, 1, 2, \dots, 98. \quad (3.5)$$

Stejně jako u longitudinálního pohledu se i zde liší vzorce (3.4) a (3.5) o pohlaví. Výhoda transversálního přístupu spočívá v analýze aktuálního rozložení intenzit úmrtnosti. Tímto postupem se ale ztrácí přesnost, protože dochází k výpočtu z odlišných generací. Tato skutečnost je ilustrována na vzorcích (3.4) i (3.5), kde v čitateli je počet žijících z generace o rok nižší než počet žijících ve jmenovateli. Transverzální přístup má své nedostatky, ale zároveň jeho zjednodušující charakter přispívá ke značné popularitě. I z tohoto důvodu bude použit v analytické části této práce.

Pravděpodobnosti přežití mužů a žen tvoří základ projekce žijících. Kučera (1998, s. 75) v tento moment zavádí tzv. krok transformace označený znakem  $k$ , který představuje, v jaké fázi se výpočet nachází. Aubrecht (2016, s. 19) tento znak nazývá projekčním krokem. Nultý krok transformace označuje výchozí bod počtu žijících, tzn. skutečnou strukturu počtu žijících. Při přechodu do prvního kroku transformace se struktura počtu žijících musí odhadnout (Pavlík a kol., 1986, s. 397; Kučera, 1998, s. 75)

$${}^{z-1-x}_{t+1}\hat{P}_{x+1}^{k+1,f} = {}^{z-1-x}_t P_x^{k,f} * {}^{z-1-x}_t S_x^f, \quad x=0,1,...,98, \quad k=0, \quad (3.6)$$

$${}^{z-1-x}_{t+1}\hat{P}_{x+1}^{k+1,m} = {}^{z-1-x}_t P_x^{k,m} * {}^{z-1-x}_t S_x^m, \quad x=0,1,...,98, \quad k=0. \quad (3.7)$$

Základní myšlenkou je, že při zvoleném výchozím okamžiku v rámci druhého hlavního souboru dochází k posunu původní věkové struktury žijících do nového roku. Na původní strukturu žijících, která je zde charakterizována skutečnými počty žijících, jsou aplikovány pravděpodobnosti přežití, které snižují počty žijících o zemřelé jedince. Výhodou výpočtu v druhém hlavním souboru je, že zachovává shodné generace. Tímto způsobem dochází k odhadu počtu žijících mužů (3.6) a žen (3.7) v dokončeném věku  $x + 1$  v roce  $t + 1$  z generace  $z - 1 - x$  v prvním kroku transformace. Odhad počtu žijících v prvním kroku transformace slouží jako podklad pro výpočet odhadu počtu žijících ve druhém a následně vyšším kroku transformace.

Obecně platí, že odhad počtu žijících ve vyšším kroku transformace se vypočítá z již odhadnutého počtu žijících v předchozím kroku transformace (Pavlík a kol., 1986, s. 397; Kučera, 1998, s. 75)

$${}^{z-1-x}_{t+1}\hat{P}_{x+1}^{k+1,f} = {}^{z-1-x}_t \hat{P}_x^{k,f} * {}^{z-1-x}_t S_x^f, \quad x=0,1,...,98, \quad k > 0, \quad (3.8)$$

$${}^{z-1-x}_{t+1}\hat{P}_{x+1}^{k+1,m} = {}^{z-1-x}_t \hat{P}_x^{k,m} * {}^{z-1-x}_t S_x^m, \quad x=0,1,...,98, \quad k > 0. \quad (3.9)$$

V rámci druhého hlavního souboru je struktura žijících postupně snižována o počty zemřelých. Ve výpočtu je tento efekt pro vyšší krok transformace vyjádřen vynásobením odhadnuté struktury žijících z předchozího kroku transformace a příslušnými pravděpodobnostmi přežití v dokončeném věku. Tím se struktura žijících posune do dalšího roku. Obecně platí, že se nejprve odhadnou počty žijících žen (3.6) a mužů (3.7) v dokončeném věku  $x + 1$  v prvním kroku transformace a tento odhad slouží jako základ odhadu počtu žijících žen (3.8) a mužů (3.9) ve druhém kroku transformace. Nově odhadnutá struktura žijících ve druhém kroku transformace je podkladem pro odhad počtu žijících žen (3.8) a mužů (3.9) ve třetím kroku transformace. Tento postup se opakuje, až se dosáhne horizontu projekce. V této fázi výpočet končí.

Druhým bodem při výpočtu kohortně komponentní metodou bez migrace je tzv. **projekce žijících v nejvyšších věcích**. Před tím, než je možné sestavit odhad počtu žijících v posledním otevřeném intervalu, se musí nejprve vypočítat pravděpodobnosti přežití v nejvyšších dokončených věcích. Tyto pravděpodobnosti podléhají zkreslení především kvůli malému počtu žijících osob a rovněž proto, že poslední otevřený interval si volí autor dle vlastního uvážení, případně podle vžitých konvencí. Řešení navrhuje Klosterman (1990, s. 75)

$${}^{z-1-x}_t S_x^f = \frac{{}_t T_x^f}{{}_t T_{x-1}^f}, \quad x=99, \quad (3.10)$$

$${}^{z-1-x}_t S_x^m = \frac{{}_t T_x^m}{{}_t T_{x-1}^m}, \quad x=99, \quad (3.11)$$

$${}^{z-1-(x+)}_t S_{x+}^f = \frac{{}_t T_{x+}^f}{{}_t T_{x-1}^f}, \quad x=100, \quad (3.12)$$

$${}^{z-1-(x+)}_t S_{x+}^m = \frac{{}_t T_{x+}^m}{{}_t T_{x-1}^m}, \quad x=100. \quad (3.13)$$

Princip výpočtu je založen na podílu celkového počtu let zbývajících k dožití tabulkové populaci v dokončeném věku  $x$  a celkového počtu let zbývajících k dožití tabulkové populaci

v dokončeném věku  $x$  sníženém o jeden rok. Tímto způsobem se vyrovnají pravděpodobnosti přežití žen (3.10) a mužů (3.11) v dokončeném věku 99 a pravděpodobnosti přežití žen (3.12) a mužů (3.13) v dokončeném věku 100 a více let. Jinou otázkou jsou metody vyrovnávání úmrtnostních tabulek a to nejen v nejvyšších věcích. Nejčastěji se vyrovnávají tabulkové míry úmrtnosti a pravděpodobnosti přežití v přesném věku. Na tyto funkce se aplikují modely, které vyhladí křivku úmrtnosti. Z nově vyrovnaných funkcí se poté přepočítá celá úmrtnostní tabulka. Vyrovnání má vliv i na pravděpodobnosti přežití v dokončeném věku. Tímto problémem se zabývá například Burcin, Tesárková a Šídlo (2010, s. 77–89).

Pravděpodobnosti přežití v dokončeném věku 99 a 100 a více let jsou podkladem pro projekci žijících v nejvyšších věcích. Odhad počtu žijících v posledním otevřeném intervalu v prvním kroku transformace se vypočítá (Klosterman, 1990, s. 76)

$${}^{z-1-(x+)}_{t+1} \hat{P}_{x+}^{k+1,f} = {}^{z-1-(x+)}_t P_{x+}^{k,f} * {}^{z-1-(x+)}_t S_{x+}^f + {}^{z-x}_t P_{x-1}^{k,f} * {}^{z-x}_t S_{x-1}^f, \quad x=100, \quad k=0, \quad (3.14)$$

$${}^{z-1-(x+)}_{t+1} \hat{P}_{x+}^{k+1,m} = {}^{z-1-(x+)}_t P_{x+}^{k,m} * {}^{z-1-(x+)}_t S_{x+}^m + {}^{z-x}_t P_{x-1}^{k,m} * {}^{z-x}_t S_{x-1}^m, \quad x=100, \quad k=0. \quad (3.15)$$

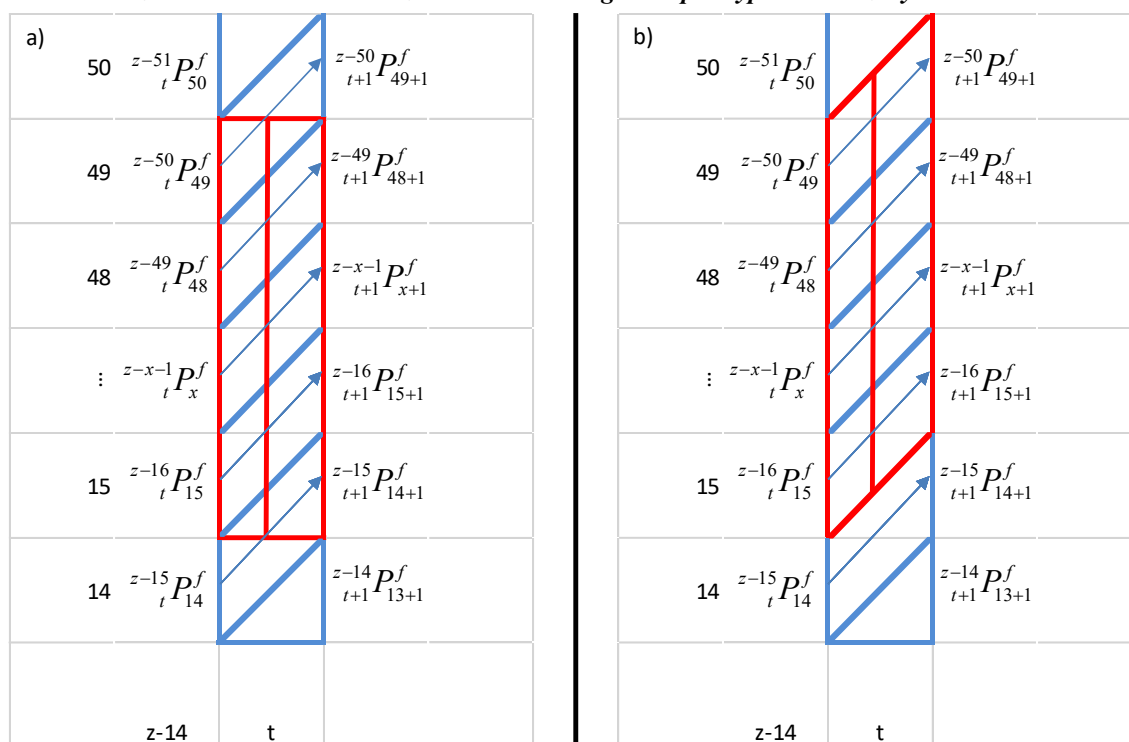
Rovnice (3.14) a (3.15) se liší o pohlaví. Nultý krok transformace představuje výpočet ze skutečného počtu žijících a přechod do dalšího kroku transformace znázorňuje odhad počtu žijících. Hlavní myšlenka v rámci druhého hlavního souboru spočívá v zachování přeživších žijících v posledním otevřeném intervalu a posunu počtu žijících v dokončeném věku 99 do posledního otevřeného intervalu.

Pro vyšší krok transformace je odhad počtu žijících v posledním otevřeném intervalu spočítán (Klosterman, 1990, s. 76)

$${}^{z-1-(x+)}_{t+1} \hat{P}_{x+}^{k+1,f} = {}^{z-1-(x+)}_t \hat{P}_{x+}^{k,f} * {}^{z-1-(x+)}_t S_{x+}^f + {}^{z-x}_t \hat{P}_{x-1}^{k,f} * {}^{z-x}_t S_{x-1}^f, \quad x=100, \quad k > 0, \quad (3.16)$$

$${}^{z-1-(x+)}_{t+1} \hat{P}_{x+}^{k+1,m} = {}^{z-1-(x+)}_t \hat{P}_{x+}^{k,m} * {}^{z-1-(x+)}_t S_{x+}^m + {}^{z-x}_t \hat{P}_{x-1}^{k,m} * {}^{z-x}_t S_{x-1}^m, \quad x=100, \quad k > 0. \quad (3.17)$$

V principu jsou vzorce pro ženy (3.16) a pro muže (3.17) totožné s rovnicemi (3.14) a (3.15). Jediný rozdíl spočívá v tom, že se počty žijících počítají ne ze skutečných počtů, ale z odhadnutých, a v zásadě navazují na již uvedené rovnice (3.14) a (3.15).

**Obr. 5: Znázornění středního stavu žen v Lexisově diagramu pro výpočet narozených**

**Zdroj:** vlastní zpracování.

Třetím a posledním bodem ve výpočtu je **projekce narozených**. Na obrázku 5 v kategorii *a* je graficky naznačen výpočet středního stavu žen ve třetím hlavním souboru pro odhad počtu narozených. Aby bylo možné sestavit projekci narozených, je zapotřebí nejen spočítat střední stav žen, ale také míry plodnosti podle věku ke střednímu stavu roku  $t$ . Ty se vypočítají pro třetí hlavní soubor (Pressat, 2008, s. 49, Pavlík a kol., 1986, s. 296)

$$f_x = \frac{\frac{z-1-x}{t} B_x^v + \frac{z-x}{t} B_x^v}{\frac{z-1-x}{t} P_x^f + \frac{z-x}{t+1} P_x^f}, \quad x=15,16,\dots,49. \quad (3.18)$$

Podstatou měř plodnosti podle věku, bez ohledu na typ hlavního souboru, je podíl počtu živě narozených ke střednímu stavu žen. Specifičnost třetího hlavního souboru spočívá v jeho rozdělení do dvou elementárních souborů, kde každý z nich pochází z jiné generace. Tento postup je možné ilustrovat na vzorci (3.18), kde v čitateli je součet dvou elementárních souborů počtu živě narozených. První elementární soubor zahrnuje živě narozené v roce  $t$  ženám, které pochází z generace  $z-l-x$  v dokončeném věku  $x$ . Druhý elementární soubor je tvořen živě narozenými v roce  $t$  ženám z generace  $z-x$  v dokončeném věku  $x$ . Ve jmenovateli se počítá střední stav z počtu žijících žen k výchozímu okamžiku v roce  $t$  z generace  $z-l-x$  v dokončeném věku  $x$  a z počtu žijících žen k výchozímu okamžiku v roce  $t+1$  z generace  $z-x$  v dokončeném věku  $x$ .

Na obrázku 5 v kategorii *b* je znázorněn výpočet středního stavu žen ve druhém hlavním souboru. Tento střední stav slouží k odhadu počtu narozených. Pro výpočet je ale zapotřebí také znát míry plodnosti z generace  $z-l-x$  ke střednímu stavu roku  $t$  spočítané z druhého hlavního souboru (Pressat, 2008, s. 50, Pavlík a kol., 1986, s. 296)

$${}_{t+\frac{1}{2}}^{z-1-x}f = \frac{{}_{t+\frac{1}{2}}^{z-1-x}B_x^v + {}_{t+\frac{1}{2}}^{z-1-x}B_{x+1}^v}{{}_{t+\frac{1}{2}}^{z-1-x}P_x^f + {}_{t+\frac{1}{2}}^{z-1-x}P_{x+1}^f}, \quad x=15,16,\dots,49. \quad (3.19)$$

Druhý hlavní soubor oproti třetímu zahrnuje stejné generace a tím je i výpočet mnohem přesnější. Míry plodnosti (3.19) představují podíl počtu živě narozených ke střednímu stavu žen. Počet narozených je charakterizován součtem dvou elementárních souborů počtu živě narozených v rámci druhého hlavního souboru. Střední stav žen vyjadřuje průměr z počtu žijících žen na počátku, případně na konci intervalu roku  $t$ , a z počtu žijících žen na počátku, případně na konci intervalu roku  $t + 1$ , při uvažování druhého hlavního souboru.

Výběr typu hlavního souboru u měř plodnosti má vliv i na strukturu výpočtu živě narozených. Pokud byly zvoleny míry plodnosti podle věku spočítané z druhého hlavního souboru, pak se odhad počtu narozených v roce  $t$  v  $k + 1$  kroku transformace vypočítá (Kučera, 1998, s. 76; Roubíček, 1996, s. 212)

$${}_t\hat{B}^{v,k+1} = \sum_{x=15}^{49} \frac{({}_{t+\frac{1}{2}}^{z-1-x}P_x^{k,f} + {}_{t+\frac{1}{2}}^{z-1-x}P_x^{k,f} * {}_{t+\frac{1}{2}}^{z-1-x}S_x^f)}{2} * {}_{t+\frac{1}{2}}^{z-1-x}f, \quad k=0, \quad (3.20)$$

$${}_t\hat{B}^{v,k+1} = \sum_{x=15}^{49} \frac{({}_{t+\frac{1}{2}}^{z-1-x}\hat{P}_x^{k,f} + {}_{t+\frac{1}{2}}^{z-1-x}\hat{P}_x^{k,f} * {}_{t+\frac{1}{2}}^{z-1-x}S_x^f)}{2} * {}_{t+\frac{1}{2}}^{z-1-x}f, \quad k>0, \quad (3.21)$$

Jádrem výpočtu (3.20) je vynásobení měř plodnosti se středním stavem žen, který je vyjádřen jako průměr skutečného počtu žijících a odhadem počtu žijících v novém roce. Tím se získá odhad počtu narozených v roce  $t$  ženám ve věku  $x$  v prvním kroku transformace. Jejich součtem, který je formulován pro věky 15 až 49 let, se odhadne celkový počet živě narozených v roce  $t$  v prvním kroku transformace. Ve vzorci (3.21) je naznačen výpočet pro vyšší krok transformace s tím, že se již počítá z odhadnutých počtů žijících žen a ne se skutečným, jako tomu bylo v předchozím případě.

Častěji bývají zvoleny míry plodnosti (3.18) spočítané z třetího hlavního souboru a z nich vychází i samotný odhad počtu narozených v roce  $t$  v  $k + 1$  kroku transformace (Kučera, 1998, s. 76; Roubíček, 1996, s. 212)

$${}_t\hat{B}^{v,k+1} = \sum_{x=15}^{49} \frac{({}_{t+\frac{1}{2}}^{z-1-x}P_x^{k,f} + {}_{t+\frac{1}{2}}^{z-x}P_{x-1}^{k,f} * {}_{t+\frac{1}{2}}^{z-x}S_{x-1}^f)}{2} * {}_{t+\frac{1}{2}}^{z-1-x}f_x, \quad k=0, \quad (3.22)$$

$${}_t\hat{B}^{v,k+1} = \sum_{x=15}^{49} \frac{({}_{t+\frac{1}{2}}^{z-1-x}\hat{P}_x^{k,f} + {}_{t+\frac{1}{2}}^{z-x}\hat{P}_{x-1}^{k,f} * {}_{t+\frac{1}{2}}^{z-x}S_{x-1}^f)}{2} * {}_{t+\frac{1}{2}}^{z-1-x}f_x, \quad k>0, \quad (3.23)$$

Celkový počet narozených v prvním kroku transformace se vypočítá podle vzorce (3.22), kdy podstata zůstává stejná jako u vzorce (3.20), tedy vynásobení středního stavu žen a měř plodnosti podle věku. Rozdíl spočívá v principu výpočtu středního stavu a měř plodnosti. Míry plodnosti podle věku byly již rozebrány ve vzorci (3.18). Střední stav žen spočítaný ze třetího hlavního souboru představuje průměr počtu žijících žen z generace  $z - 1 - x$  a odhadu počtu žijících žen z generace  $z - x$ . Oba počty žijících žen pocházejí z jiné generace a to způsobuje zkreslení. Vzorec (3.23) je určen pro vyšší krok transformace, kdy výpočet středního stavu se nepočítá ze skutečného počtu žijících žen, nýbrž z odhadnutého z předchozího kroku. Jelikož v aplikované části jsou k dispozici míry plodnosti podle věku spočítané ze třetího hlavního souboru, je použit právě tento postup.

Dalším úkolem je stanovit z celkového počtu narozených dětí, kolik se narodilo chlapců a kolik dívek. To lze určit z následujících rovnic (Kučera, 1998, s. 77)

$${}_t\hat{B}^{v,k+1,f} = \phi * {}_t\hat{B}^{v,k+1}, \quad k \geq 0, \quad (3.24)$$

$${}_t\hat{B}^{v,k+1,m} = (1 - \phi) * {}_t\hat{B}^{v,k+1}, \quad k \geq 0. \quad (3.25)$$

Odhad počtu narozených dívek (3.24) v roce  $t$  v  $k + 1$  kroku transformace se vypočítá vynásobením podílu děvčat při narození, který je vyjádřen řeckým písmenkem  $\phi$  a odhadnutým celkovým počtem narozených v roce  $t$  v  $k + 1$  kroku transformace. Das Gupta a kol. (2003, s. 156) ve svém článku uvádějí, že poměr pohlaví při narození se pohybuje 105–106 narozených chlapců na 100 dívek. Pro poměr 106 chlapců na 100 dívek se podíl děvčat při narození pohybuje přibližně na hodnotě 0,485. Tato hodnota  $\phi$  bude použita v analytické části této práce. Odhad počtu narozených chlapců (3.25) v roce  $t$  v  $k + 1$  kroku transformace se vypočítá podobně jako odhad počtu narozených dívek, pouze se na celkový počet narozených dětí aplikuje podíl chlapců při narození charakterizovaný jako  $1 - \phi$ . Přesněji lze podíl děvčat při narození v roce  $t$  formulovat vzorcem inspirovaným knihou Pavlíka a kol. (1986, s. 110)

$${}_t\phi = \frac{{}_tB^{v,f}}{{}_tB^v}. \quad (3.26)$$

Výpočet (3.26) zahrnuje podíl živě narozených dívek v roce  $t$  vztažený k celkovému počtu živě narozených v roce  $t$  (Kučera, 1998, s. 77). Do budoucna by bylo nutné tento podíl odhadnout.

Posledním krokem při řešení kohortně komponentní metody bez migrace je posun živě narozených do výchozího okamžiku roku  $t + 1$ . Pokud je uvažován **longitudinální přístup** výpočet lze napsat (Shryock a kol., 1973, s. 816; Kučera, 1998, s. 77)

$${}_{t+1}^z\hat{P}_0^{k+1,f} = {}_t\hat{B}^{v,k+1,f} * \frac{{}_tL_0^f}{{}_tI_0^f}, \quad k \geq 0, \quad (3.27)$$

$${}_{t+1}^z\hat{P}_0^{k+1,m} = {}_t\hat{B}^{v,k+1,m} * \frac{{}_tL_0^m}{{}_tI_0^m}, \quad k \geq 0. \quad (3.28)$$

Základní myšlenka spočívá v posunu odhadu narozených v roce  $t$  do výchozího okamžiku počtu žijících žen (3.27) a mužů (3.28) v dalším roce. Při tomto posunu dochází k redukci narozených o zemřelé jedince. Ve výpočtu je tento efekt charakterizován vynásobením odhadu narozených s podílem tabulkového počtu žijících žen v dokončeném věku nula a tabulkovým počtem dožívajících se přesného věku nula viz obrázek 4. Jelikož je uvažován longitudinální přístup, celý výpočet se realizuje v rámci stejné generace. K tomuto postupu jsou zapotřebí longitudinální úmrtnostní tabulky. Odhady počtu narozených se vypočítají ze vzorce pro ženy (3.20) a muže (3.21).

Druhým, a dá se říci i častěji používaným postupem, je **transverzální přístup**, který nezohledňuje generace. Posun živě narozených do výchozího okamžiku roku  $t + 1$  lze znázornit (Shryock a kol., 1973, s. 816; Kučera, 1998, s. 77)

$${}_{t+1}^z\hat{P}_0^{k+1,f} = {}_t\hat{B}^{v,k+1,f} * \frac{{}_tL_0^f}{{}_tI_0^f}, \quad k \geq 0, \quad (3.29)$$

$${}_{t+1}^z\hat{P}_0^{k+1,m} = {}_t\hat{B}^{v,k+1,m} * \frac{{}_tL_0^m}{{}_tI_0^m}, \quad k \geq 0. \quad (3.30)$$

Princip výpočtu zůstává stejný jako u longitudinálního přístupu, kdy se vynásobí odhad počtu narozených s podílem tabulkového počtu žijících žen v dokončeném věku nula a kořenem tabulky. Tím se získá odhad počtu žijících žen (3.29) a mužů (3.30) v dokončeném věku nula v roce  $t + I$  v  $k + I$  kroku transformace z generace  $z$ . Odhady počtu narozených jsou zpracovány z třetího hlavního souboru pro ženy z rovnice (3.22) a pro muže z výpočtu (3.23).

### 3.1.2. Kohortně komponentní metoda s migrací

Kohortně komponentní metoda s migrací oproti kohortně komponentní metodě bez migrace rozšiřuje odhad budoucího vývoje o komponentu migrace. Je založena na bilanční rovnici (Sadalla, 2005, s. 263)

$$P_t = P_0 + B - D + I - E. \quad (3.31)$$

Celkový počet žijících obyvatel (3.31) na konci roku  $t$  označený  $P_t$  se vypočítá jako celkový počet žijících k danému výchozímu stavu pod zkratkou  $P_0$ , ke kterému se od výchozího okamžiku do roku  $t$  přičtou narození  $B$ , odečtou zemřelí  $D$ , přičtou přistěhovalí  $I$  a odečtou vystěhovalí  $E$ . Stejně jako u kohortně komponentní metody bez migrace i zde je výchozí okamžik označen nulou.

Hoque a kol. (2013, s. 12) zjednodušují předchozí bilanční rovnici zavedením tzv. čisté migrace

$$P_t = P_0 + B - D + NM. \quad (3.32)$$

Pavlík a kol. (1986, s. 486) užívají pro pojem čistá migrace také synonyma migrační saldo nebo migrační přírůstek. Bilanční rovnice (3.32) se liší od vzorce (3.31) právě v pojetí migračního salda, které se vypočítá (Roubíček, 1996, s. 169)

$$NM = I - E. \quad (3.33)$$

Migrační saldo (3.33) již ve své podstatě zahrnuje rozdíl přistěhovalých a vystěhovalých. Převážně pro menší územní celky publikují statistické úřady často pouze migrační saldo a nikoli počty přistěhovalých a vystěhovalých. Z tohoto důvodu byla zařazena rovnice (3.32).

Kučera (1998, s. 81) ve své práci projektuje migraci po složkách. Zahrnuje jak imigrační, tak i emigrační složku. K tomuto přístupu jsou zapotřebí (Kučera, 1998, s. 85):

- Míry vystěhování podle věku a pohlaví.
- Počty přistěhovalých podle věku a pohlaví.

Pavlík a kol. (1986, s. 402) tvrdí, že přesnějším postupem je zahrnutí vystěhovalých a přistěhovalých do projekce obyvatel, ale pro praxi je postačující uvažování migračního salda. Překážkou je především nedostatek dat. K realizaci tohoto přístupu jsou nezbytné následující údaje:

- Migrační saldo podle věku a pohlaví.

Existují také situace, kdy nejsou k dispozici ani tyto údaje. V tomto případě se může přistoupit k odhadu migračního salda. Více o tomto přístupu, který bude použit v analytické části této práce, v podkapitole 3.4.

Základním problémem kohortně komponentní metody s migrací je navržení způsobu zakomponování migrace. Pavlík a kol. (1986, s. 402) navrhuje dva způsoby zahrnutí migrace. První z nich zohledňuje postupné zakomponování migračního salda, tzn. předpokládá se, že i migranti realizují svou plodnost a také umírají. Druhý způsob modifikuje konečné počty žijících



o odhadované migrační saldo. V tabulce 2 je uveden seznam symbolů, které budou užity v následujícím textu.

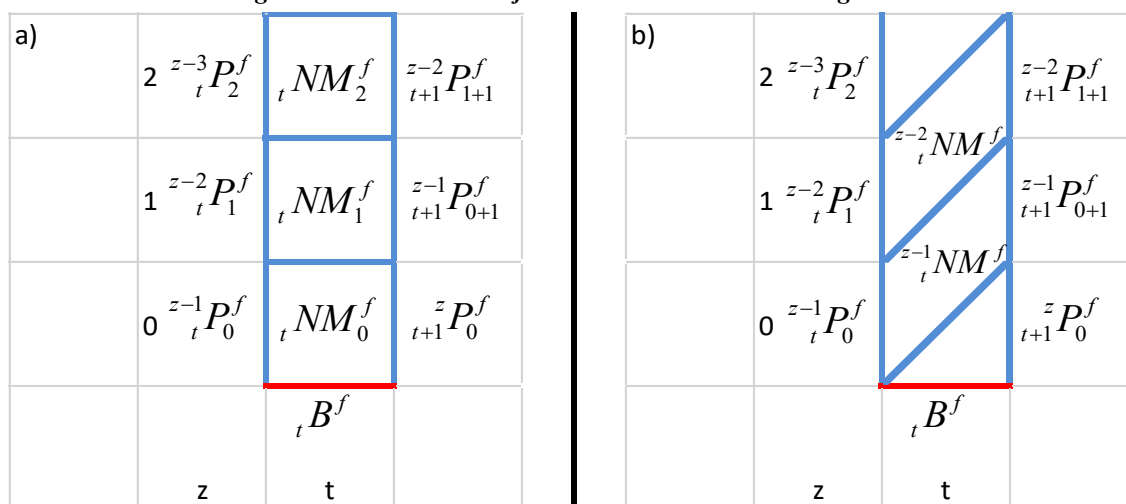
**Tab. 2: Seznam symbolů použitých u kohortně komponentní metody s migrací**

Symbol	Popis
$f$	ženy
$m$	muži
$x$	dokončený věk
$t$	rok
$z$	generace
$k$	krok transformace
${}^{z-1-x}_t \widehat{NM}^{k+1,f}$	odhad migrační saldo žen z generace $z - l - x$ v roce $t$ v $k + 1$ kroku transformace spočítané z druhého hlavního souboru
${}_t \widehat{NM}_x^{k+1,f}$	odhad migrační saldo žen v dokončeném věku $x$ v roce $t$ v $k + 1$ kroku transformace spočítané z třetího hlavního souboru
${}^{z-1-x}_t P_x^{k,f}$	počet žijících žen v dokončeném věku $x$ k výchozímu okamžiku v roce $t$ z generace $z - l - x$ v $k$ -tém kroku transformace
${}^{z-1-x}_{t+1} P_{x+1}^{k,f*}$	počet žijících žen se zakomponováním migrace v dokončeném věku $x + l$ k výchozímu okamžiku v roce $t + l$ z generace $z - l - x$ v $k$ -tém kroku transformace
${}^{z-1-x}_t \hat{P}_x^{k,f}$	odhad počtu žijících žen v dokončeném věku $x$ k výchozímu okamžiku v roce $t$ z generace $z - l - x$ v $k$ -tém kroku transformace
${}^{z-1-x}_{t+1} \hat{P}_{x+1}^{k,f*}$	odhad počtu žijících žen se zakomponováním migrace v dokončeném věku $x + l$ k výchozímu okamžiku v roce $t + l$ z generace $z - l - x$ v $k$ -tém kroku transformace

**Zdroj:** Kučera (1998, s. 75–78), Hoque a kol. (2013, s. 12), vlastní zpracování.

**První způsob zahrnutí migrace** je podle Pavlíka a kol (1986, s. 402) komplikovanější, ale zároveň poskytuje přesnější výsledky. Preston a kol. (2001, s. 127) ho zjednodušují do dvou etap, kde předpokladem užití je rovnoměrné rozložení migrace v celém intervalu. V první etapě se přiřadí polovina migračního salda k počtu žijících na počátku intervalu. Tímto způsobem upravené počty žijících slouží jako nový vstup do kohortně komponentní metody bez migrace. Druhá etapa spočívá v zařazení zbylé poloviny migračního salda ke konci intervalu nově odhadnutého počtu žijících. Přístup k oběma etapám je zvolen podle toho, z jakého hlavního souboru je vypočítáno migrační saldo. Uvažován bude druhý a třetí hlavní soubor.

**Obr. 6: Znázornění migračního salda žen v nejnižších věcích v Lexisově diagramu**



**Zdroj:** vlastní zpracování.

Obrázek 6 je rozdělen do dvou kategorií. Kategorie *a* zahrnuje migrační saldo žen v nejnižších věcích, které je vypočítané v rámci třetího hlavního souboru. Kategorie *b* znázorňuje migrační saldo žen v nejnižších věcích spočítané prostřednictvím druhého hlavního souboru. V obou kategoriích jsou pro názornost grafického zobrazení výpočtu kohortně komponentní metody s migrací doplněny i počty žijících *P* a počty narozených *B*.

**Druhý hlavní soubor** zohledňuje přesnější postup zahrnutí migrace. Jeho úskalí spočívá v nedostatku dat. Často není k dispozici migrační saldo členěné do třetího hlavního souboru a pouze výjimečně do druhého hlavního souboru. Přesto existuje způsob, jakým lze odhadnout migrační saldo v druhém hlavním souboru. Jak již bylo řečeno, tento způsob bude popsán v podkapitole 3.4.

V první etapě výpočtu kohortně komponentní metody s migrací se formulují počty žijících se zakomponováním migrace v nultém kroku transformace v rámci druhého hlavního souboru (McCAA, 1989, s. 294)

$${}^{z-1-x}P_x^{k,f*} = {}^{z-1-x}P_x^{k,f} + \frac{{}_t\widehat{NM}^{k+1,f}}{2}, \quad x = 0, 1, \dots, 100+, \quad k = 0, \quad (3.34)$$

$${}^{z-1-x}P_x^{k,m*} = {}^{z-1-x}P_x^{k,m} + \frac{{}_t\widehat{NM}^{k+1,m}}{2}, \quad x = 0, 1, \dots, 100+, \quad k = 0. \quad (3.35)$$

Princip této transformace spočívá v přiřazení poloviny migračního salda ke skutečnému počtu žijících. Tím se vypočítají skutečné počty žijících žen (3.34) a mužů (3.35) se zohledněním migrace. Nově zpracované počty žijících slouží jako podklad do kohortně komponentní metody bez migrace. Postupně se vypočítá projekce žijících charakterizovaná rovnicemi (3.6) a (3.7), projekce žijících v nejvyšších věcích popsána vzorci (3.14) a (3.15) a nakonec i projekce narozených. Predikce počtu narozených se vypočítá v závislosti na typu souboru. Pro druhý hlavní soubor se odhadnou narození z formulace (3.20) a pro třetí hlavní soubor z rovnice (3.22). Následně se odhadnou počty narozených chlapců a dívek a projektují se na konec intervalu.

Druhá etapa zahrnuje zbylou část migračního salda, která se přiřadí k odhadnutému počtu žijících v dokončeném věku  $x + 1$  v roce  $t + 1$  z generace  $z - 1 - x$  nejprve v nultém kroku transformace (McCAA, 1989, s. 294)

$${}^{z-1-x}\hat{P}_{x+1}^{k+1,f*} = {}^{z-1-x}\hat{P}_{x+1}^{k+1,f} + \frac{{}_t\widehat{NM}^{k+1,f}}{2}, \quad x = 1, 2, \dots, 98, \quad k \geq 0, \quad (3.36)$$

$${}^{z-1-x}\hat{P}_{x+1}^{k+1,m*} = {}^{z-1-x}\hat{P}_{x+1}^{k+1,m} + \frac{{}_t\widehat{NM}^{k+1,m}}{2}, \quad x = 1, 2, \dots, 98, \quad k \geq 0. \quad (3.37)$$

Vyšší krok transformace bude popsán později s užitím stejného vzorce. Základem výpočtu je odhad počtu žijících z první etapy a nultého kroku transformace, k té se přičte druhá polovina migračního salda vymezeného druhým hlavním souborem. Tím se získá odhad počtu žijících žen (3.36) a mužů (3.37) se zakomponováním migrace. Předpokládá se, že na druhou polovinu migračního salda nepůsobí úmrtnostní vlivy ani plodnost. Rovnice (3.36) i (3.37) jsou definovány pro věky 1 až 98.

Otázkou je, jakým způsobem zahrnout migraci v posledním otevřeném intervalu u druhého hlavního souboru. Tento problém řeší rovnice (Preston a kol., 2001, s. 128)

$${}^{z-(x+)}\hat{P}_{x+}^{k+1,f*} = {}^{z-(x+)}\hat{P}_{x+}^{k+1,f} + \frac{{}_t\widehat{NM}^{k+1,f}}{2} + \frac{{}_t\widehat{NM}^{k+1,f}}{2}, \quad x = 100, \quad k \geq 0, \quad (3.38)$$

$${}^{z-(x+)}_{t+1}\hat{P}_{x+}^{k+1,m*} = {}^{z-(x+)}_{t+1}\hat{P}_{x+}^{k+1,m} + \frac{{}^{z-x}\widehat{NM}_x^{k+1,m}}{2} + \frac{{}^{z-1-(x+)}_t\widehat{NM}_x^{k+1,m}}{2}, \quad x=100, \quad k \geq 0. \quad (3.39)$$

Hlavním smyslem je přiřazení poloviny migračního salda z generace  $z - 100$  a poloviny migračního salda z posledního otevřeného intervalu k odhadu počtu žijících. Obecně se do posledního otevřeného intervalu dostávají dvě generace. Tím se získá odhad počtu žijících žen (3.38) a mužů (3.39) v posledním otevřeném intervalu se zakomponováním migrace v prvním kroku transformace.

Odlišnost výpočtu také nastává v nejnižším věku, kdy se živě narození přesouvají do výchozího okamžiku nového roku. Navazuje se na výpočet v závislosti na longitudinálním pohledu (3.27) a (3.28) nebo transversálním pohledu (3.29) a (3.30). Migrační saldo zahrnuté v dolním elementárním souboru z generace  $z$  se zahrne (Preston a kol., 2001, s. 128)

$${}^{z}_{t+1}\hat{P}_x^{k+1,f*} = {}^{z}_{t+1}\hat{P}_x^{k+1,f} + {}^z\widehat{NM}_x^{k+1,f}, \quad x=0, \quad k \geq 0, \quad (3.40)$$

$${}^{z}_{t+1}\hat{P}_x^{k+1,m*} = {}^{z}_{t+1}\hat{P}_x^{k+1,m} + {}^z\widehat{NM}_x^{k+1,m}, \quad x=0, \quad k \geq 0. \quad (3.41)$$

Základem této transformace je přiřazení migračního salda z generace  $z$  k odhadnutému počtu žijících v roce  $t + 1$ . Tím se získají odhadnuté počty žijících žen (3.40) a mužů (3.41) v dokončeném věku nula v roce  $t + 1$  v prvním kroku transformace se zohledněním migrace. Nově vzniklé odhady slouží jako podklad pro výpočet ve vyšším kroku transformace.

Pro vyšší krok transformace se začíná znovu u první etapy výpočtu. Odhad počtu žijících se zakomponováním migrace ve vyšším kroku transformace se vypočítá (McCAA, 1989, s. 294)

$${}^{z-1-x}_t\hat{P}_x^{k,f*} = {}^{z-1-x}_t\hat{P}_x^{k,f} + \frac{{}^{z-1-x}\widehat{NM}_x^{k+1,f}}{2}, \quad x=0,1,\dots,100+, \quad k > 0, \quad (3.42)$$

$${}^{z-1-x}_t\hat{P}_x^{k,m*} = {}^{z-1-x}_t\hat{P}_x^{k,m} + \frac{{}^{z-1-x}\widehat{NM}_x^{k+1,m}}{2}, \quad x=0,1,\dots,100+, \quad k > 0. \quad (3.43)$$

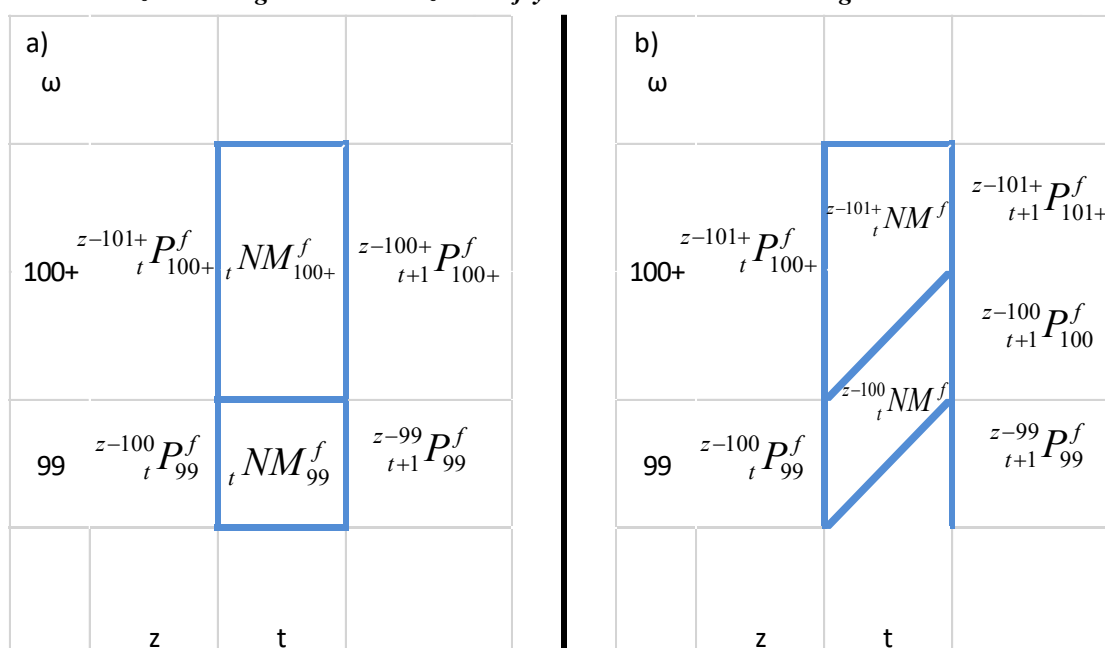
Rozdíl mezi rovnicemi (3.42) a (3.34) pro ženy a vzorci (3.43) a (3.35) pro muže spočívá v tom, že počty žijících se zakomponováním migrace ve vyšším kroku transformace se spočítají z odhadnutého počtu žijících na rozdíl od nultého kroku transformace počítaného ze skutečného počtu žijících. Přičemž smysl zůstává stejný, a to přiřazení poloviny migračního salda v rámci druhého hlavního souboru k odhadu počtu žijících. Tímto způsobem vzniklý odhad se zakomponováním migrace slouží jako vstup do kohortně komponentní metody bez migrace. Postupně se ve vyšším kroku transformace vypočítá projekce žijících žen (3.8) a mužů (3.9), projekce žijících žen (3.16) a mužů (3.17) v posledním otevřeném intervalu a projekce narozených pro druhý hlavní soubor ze vzorce (3.21) nebo pro třetí hlavní soubor z formulace (3.23). Následně se definují počty narozených chlapců a dívek a stanoví se počet těch, kteří přežili k výchozímu okamžiku nového roku.

Nově vzniklé odhady z celé projekce žijících přecházejí do druhé etapy výpočtu pro vyšší krok transformace. K takto vzniklému odhadu se přiřadí polovina migračního salda v rámci druhého hlavního souboru. Samotný výpočet vychází ze vzorce pro ženy (3.36) a pro muže (3.37). Pro poslední otevřený interval je rovnice odlišná. Zahrnuje přeživší jedince z generace  $z - 100$  a z generace posledního otevřeného intervalu. Výpočet ve vyšším kroku transformace je charakterizován rovnicemi pro ženy (3.38) a pro muže (3.39). Důležité je také zahrnutí migračního salda z generace  $z$  v dolním elementárním souboru, které se přičte k počtu žijícím

v dokončeném věku nula v roce  $t + 1$ . Tímto způsobem se vypočítají počty žijících žen (3.40) a mužů (3.41) v dokončeném věku nula v roce  $t + 1$  ve vyšším kroku transformace se zakomponováním migrace.

Shrne-li se celý postup, pak nultý krok transformace představuje skutečné počty žijících. Prvním krokem transformace se odhadnou počty žijících včetně zakomponování migrace. Tento odhad slouží jako podklad pro druhý krok transformace. Druhý krok transformace je základem třetího a následně vyššího kroku transformace. Výpočet končí, jakmile se dosáhne horizontu projekce.

**Obr. 7: Znázornění migračního salda žen v nejvyšších věcích v Lexisově diagramu**



**Zdroj:** vlastní zpracování.

Obrázek 7 slouží jako grafický prvek pro výpočet kohortně komponentní metody s migrací u žen v posledním otevřeném intervalu. Přístup se liší v závislosti na typu hlavního souboru. V kategorii *a* je Lexisův diagram rozdělen do třetích hlavních souborů. Migrační saldo žen ve věku 100 a více let se rozdělí na polovinu. První polovina se přiřadí k počtu žijících žen na počátku intervalu věku 100 a více let a druhá polovina se přiřadí k počtu žijících žen na konec posledního otevřeného intervalu. V kategorii *b* se Lexisův diagram rozpadá do druhých hlavních souborů. Výpočet v rámci těchto souborů je složitější. Migrační saldo žen z generace charakteristické otevřeným intervalem se nejprve rozdělí na polovinu. První polovina je přičtena k počátku intervalu počtu žijících žen v dokončeném věku 100+. Tím je výpočet stejný jako u třetího hlavního souboru. Rozdíl nastává v přiřazení druhé poloviny migračního salda žen na konec intervalu počtu žijících. Jak je schematicky naznačeno v obrázku, do počtu žijících žen ve věku 100+ vstupují dvě generace migračních sald žen. Jednak generace  $z - 100$  a zároveň z generace posledního otevřeného intervalu. Obě migrační salda žen přispívají do počtu žijících jednou polovinou. Řecké písmeno omega ( $\omega$ ) v kategorii *a* i v kategorii *b* označuje věk, kterého se nikdo nedožije.

Následující výpočty budou předpokládat **třetí hlavní soubor**. Tento postup je navržen proto, že nejčastěji je migrační saldo zpracované v rámci třetího hlavního souboru. V první etapě se vypočítají počty žijících se zahrnutím migrace v nultém kroku transformace (McCAA, 1989, s. 294)

$${}^{z-1-x}_t P_x^{k,f*} = {}^{z-1-x}_t P_x^{k,f} + \frac{{}_t \widehat{NM}_x^{k+1,f}}{2}, \quad x = 0, 1, \dots, 100+, \quad k = 0, \quad (3.44)$$

$${}^{z-1-x}_t P_x^{k,m*} = {}^{z-1-x}_t P_x^{k,m} + \frac{{}_t \widehat{NM}_x^{k+1,m}}{2}, \quad x = 0, 1, \dots, 100+, \quad k = 0. \quad (3.45)$$

Rovnice (3.44) a (3.45) se liší pohlavím. Smyslem zařazení poloviny migračního salda je, že migranti umírají a realizují plodnost i během projekčního kroku. Vymírání je zohledněno v projekci žijících a v projekci žijících v nejvyšších věcích. Plodnost migrantek se realizuje právě přes polovinu migračního salda žen přičtené k celkovému počtu žijících žen. Dále se postupuje stejným způsobem jako u výpočtu z druhého hlavního souboru, a to zahrnutím počtu žijících se zakomponováním migrace do kohortně komponentní metody bez migrace.

Ve vyšším kroku transformace se odhad počtu žijících se zahrnutím migračního salda prostřednictvím třetího hlavního souboru vypočítá (McCAA, 1989, s. 294)

$${}^{z-1-x}_t \hat{P}_x^{k,f*} = {}^{z-1-x}_t \hat{P}_x^{k,f} + \frac{{}_t \widehat{NM}_x^{k+1,f}}{2}, \quad x = 0, 1, \dots, 100+, \quad k > 0, \quad (3.46)$$

$${}^{z-1-x}_t \hat{P}_x^{k,m*} = {}^{z-1-x}_t \hat{P}_x^{k,m} + \frac{{}_t \widehat{NM}_x^{k+1,m}}{2}, \quad x = 0, 1, \dots, 100+, \quad k > 0. \quad (3.47)$$

Vzorce (3.44) a (3.46) pro ženy a rovnice (3.45) a (3.47) pro muže se liší v pojetí počtu žijících. Zatímco pro nultý krok transformace představují žijící skutečné počty, pro vyšší krok transformace tvoří odhady. Nově vzniklé odhady ve vyšším kroku transformace reprezentují vstup do kohortně komponentní metody bez migrace.

Cílem druhé etapy výpočtu je zakomponování druhé poloviny migračního salda do odhadu počtu žijících (McCAA, 1989, s. 294)

$${}^{z-x}_{t+1} \hat{P}_x^{k+1,f*} = {}^{z-x}_{t+1} \hat{P}_x^{k+1,f} + \frac{{}_t \widehat{NM}_x^{k+1,f}}{2}, \quad x = 0, 1, \dots, 100+, \quad k \geq 0, \quad (3.48)$$

$${}^{z-x}_{t+1} \hat{P}_x^{k+1,m*} = {}^{z-x}_{t+1} \hat{P}_x^{k+1,m} + \frac{{}_t \widehat{NM}_x^{k+1,m}}{2}, \quad x = 0, 1, \dots, 100+, \quad k \geq 0. \quad (3.49)$$

Nejprve se navazuje na nultý krok transformace ze vzorců (3.44) a (3.45). Po zahrnutí tohoto výpočtu do kohortně komponentní metody bez migrace se získají odhady počtu žijících, které se upraví o polovinu migračního salda, viz rovnice (3.48) a (3.49). Tímto krokem se získá konečný odhad počtu žijících se zahrnutím migrace v prvním kroku transformace. Pro vyšší krok transformace slouží v první etapě rovnice (3.46) a (3.47). V druhé etapě poté vzorce (3.48) a (3.49). Výpočet končí, když se dosáhne horizontu projekce.

Shrnutím dosavadního postupu byl prezentován první způsob zahrnutí migrace, kdy se do počtu žijících zahrnuje ve dvou etapách migrační saldo v závislosti na typu hlavního souboru. Preston a kol. (2001, s. 127) tvrdí, že zjednodušení na uvažování samotného migračního salda přináší řadu problémů. Jedním z nich je, že emigrace a imigrace nemusí zahrnovat shodné příčiny. I z tohoto důvodu Kučera (1998, s. 81) tento postup dále rozšiřuje o zahrnutí imigrace a emigrace.

Pavlík a kol. (1986, s. 402) navrhuje další zpřesnění, které spočívá v uvažování odlišné plodnosti imigrantek. Velkým problémem je ale nedostatek dat. Metodiku zakomponování imigrace a emigrace do kohortně komponentní metody s migrací popisuje například Carmichael (2016, s. 377–378) nebo Preston a kol. (2001, s. 124–127). Na základě jejich poznatků bude tento postup aplikován v analytické části této práce.

**Druhý způsob zahrnutí migrace** představují Pavlík a kol. (1986, s. 402). V principu se jedná o korekci konečného počtu žijících o odhadované migrační saldo. Rozdíl od průběžného zahrnování migračního salda spočívá v tom, že migrace nevstupuje do kohortně komponentní metody průběžně a tím není uvažována její úmrtnost ani reprodukční chování. Kvůli této skutečnosti je na odhad migračního salda kladen vysoký nárok. Odhad počtu žijících se zahrnutím migrace se vypočítá z kohortně komponentní metody bez migrace a konečný odhad se upraví o migrační saldo. Podrobněji se celému postupu věnuje Klosterman (1990, s. 83) a Pavlík a kol. (1986, s. 402).

### 3.2 Vybrané modely při studiu očekávaného vývoje úmrtnosti

Podkladem pro modelování úmrtnosti jsou úmrtnostní tabulky.<sup>1</sup> V minulosti byla data o úmrtnosti hodně nespolehlivá a především existovalo malé množství dat. Tím byla vypovídající schopnost úmrtnostních tabulek velmi nízká. Docházelo k hledání zákonitostí, které by dokázaly tento problém vyřešit. Jako jeden z prvních již v roce 1725 modeloval tabulkovou funkci dožívajících Abraham De Moivre (Smith a kol., 2013, s. 225)

$$l_x = l_0 \left( 1 - \frac{x}{\omega} \right), \quad x = 12, 13, \dots, 86, \quad \omega = 86. \quad (3.50)$$

Hlavním smyslem výpočtu (3.50) je, že se od kořene tabulky odečítá sklon ke kořenu tabulky. Tento sklon je definován podílem věku a maximálního věku, kterého se již nikdo nedožije. Tím se vypočítají tabulkové počty dožívajících přesného věku  $x$ .

Významný pokrok v analytickém vyrovnání tabulkových měr úmrtnosti učinil v roce 1825 Benjamin Gompertz, který navrhl funkci o dvou parametrech, kde jeden parametr je nezávislý na věku a druhý je na věku závislý (Pavlík a kol., 1986, s. 174). William Makeham v roce 1860 rozšířil Gompertzovu myšlenku a zařadil navíc parametr, který souvisí s prostředím (Smith a kol., 2013, s. 226)

$$\mu_x = A + Bc^x, \quad x = 30, \dots, 80. \quad (3.51)$$

Gompertz-Makehamova funkce se užívá pro modelování úmrtnosti v nejvyšších věcích. Tabulkové míry úmrtnosti podle věku (3.51) zahrnují tři parametry. Parametr  $A$  je backgroundová složka, parametr  $B$  vyjadřuje hladinu úmrtnosti a parametr  $c$  značí nárůst úmrtnosti s věkem  $x$ . Burcin, Tesárková a Šídlo (2010, s. 78) tvrdí, že Gompertz-Makehamova funkce je vhodná

---

<sup>1</sup>Úmrtnostní tabulky lze vypočítat podle metodiky Prestona a kol. (2001, s. 38–50) nebo je lze převzít ze statistických úřadů.

přibližně do věku 80 let. Její dolní hranice je podle Hindeho (1998, s. 76) dobře popsána od věku 30 let.

Aby bylo možné odhadnout parametry Gompertzovy funkce, musí se nejprve tato funkce logaritmičsky transformovat. Poté je možné odhadnout parametry například metodou nejmenších čtverců. Obecně je odhad parametrů Gompertz-Makehamovy funkce složitější. Jeden ze způsobů odhadu jejích parametrů navrhuje Polášková (2013, s. 17–22) a podrobně ho metodicky zpracovává. Odhadnuté parametry se poté zpětně dosadí do rovnice (3.51) a tím se rozkolísané tabulkové míry úmrtnosti v nejvyšších věcích vyrovnají. „(...) *Parametrická vyjádření jsou využitelná v populačních projekcích díky analýze historických trendů vývoje uvažovaných parametrů* (...)“ Gavrilová a Gavrilov (2011, s. 110). K tomuto účelu budou analytické funkce pro vyrovnání úmrtnostních tabulek užity i v této práci.

Gompertzova a Gompertz-Makehamova funkce se používá k vyrovnání tabulkové míry úmrtnosti v nejvyšších věcích, případně k projekci historických parametrů v těchto nejvyšších věcích. Pro vyrovnání úmrtnostních tabulek v každém věku vytvořili v roce 1980 **Heligman a Pollard** (1980, s. 60) model

$$q_x = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + \frac{GH^x}{1 + GH^x}, \quad x = 1, 2, \dots, 99. \quad (3.52)$$

Pravděpodobnosti úmrtí (3.52) podle věku se skládají ze tří komponent. První komponenta představuje pokles úmrtnosti v dětství a reprezentují ji tři parametry. Parametr  $A$  měří úroveň úmrtnosti. Parametr  $B$  udává posun  $q_0$  oproti úrovni úmrtnosti. Čím vyšší je tato hodnota  $B$ , tím více odpovídá  $q_0$  úrovni úmrtnosti a rozdíl mezi  $q_0$  a  $q_1$  se snižuje. Parametr  $C$  reprezentuje tempo snižování úmrtnosti v dětství (Heligman a Pollard, 1980, s. 49–50). Druhá komponenta znázorňuje zvýšení úmrtnosti na vnější úmrtnost jako například nehodovost. Historicky toto zvýšení představovalo mateřskou úmrtnost kvůli těhotenství a porodu. Druhá komponenta se rovněž skládá ze tří parametrů. Prvním z nich je parametr  $D$ , který zahrnuje zvýšení úmrtnosti kvůli vnějším příčinám. Parametr  $E$  charakterizuje šířku věkového intervalu, ve kterém má ještě zvýšená úmrtnost na vnější příčiny účinnost. Parametr  $F$  značí pozici úmrtnosti na vnější příčiny. Nejčastěji se pohybuje na úrovni kolem 20 let. Třetí komponenta představuje prostřednictvím dvou parametrů úmrtnost ve stáří. Parametr  $G$  je základní složkou stárnutí a znázorňuje úroveň opotřebení lidského organismu. Parametr  $H$  značí strmost neboli sklon křivky. Posledním parametrem v modelu je věk  $x$  (Heligman a Pollard, 1980, s. 58–59). Jednotlivé komponenty jsou prezentovány v příloze 1. Další parametrické funkce popisují například Smith a kol. (2013, s. 225–230) nebo Caselli (2006, s. 177–180).

V roce 1992 publikovali Lee a Carter (1992, s. 659–671) extrapolační metodu, kterou vypočítali prognózu úmrtnosti Spojených států amerických. Jedná se o pravděpodobnostní prognózu charakteristickou intervaly spolehlivosti s určitou hladinou spolehlivosti  $(1-\alpha)$ . Používá se pro dlouhodobé odhady budoucího vývoje. Je součástí projekcí Populační divize OSN či prognóz U. S. Census Bureau (Habartová, 2009, s. 173). **Lee-Carterův model úmrtnosti** je vyjádřen vzorcem (Lee a Carter, 1992, s. 661)

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}, \quad x = 0, 1, \dots, 100+. \quad (3.53)$$

Z rovnice (3.53) představují  $m_{x,t}$  míry úmrtnosti podle věku  $x$  v roce  $t$ ,  $k_t$  označuje index úrovně úmrtnosti v roce  $t$  a symbolizuje vývoj zlogaritmovaných měr úmrtnosti v čase,  $a_x$  určuje průměrný tvar zlogaritmovaných měr úmrtnosti podle věku  $x$ ,  $b_x$  zahrnuje věkově specifické změny v úmrtnosti, které detekují citlivost úmrtnosti ve věku  $x$  na index úrovně úmrtnosti v roce  $t$  a  $\varepsilon_{x,t}$  značí chybovou složku, nezávislou na věku  $x$  a roce  $t$  (Lee a Carter, 1992, s. 660; Brouhns a kol., 2002, s. 377; Habartová, 2009, s. 174). Aby bylo možné získat jednoznačné řešení, uvádí Lee a Carter (1992, s. 661) podmínky

$$\sum_{x=0}^{100+} b_x = 1, \quad (3.54)$$

$$\sum_{t=t_0}^{t_n} k_t = 0. \quad (3.55)$$

Omezení (3.54) je konstruováno tak, že se součet věkově specifických změn v úmrtnosti rovná jedné. Druhou podmínkou (3.55) je, že se součet indexů úmrtnosti rovná nule, kde součet je definován od počátku časového období  $t_0$  po jeho konec  $t_n$ .

Při odhadu parametrů Lee-Carterova modelu se nejprve vypočítá průměrný tvar úmrtnosti podle věku (Skřivanová, 2016, s. 13)

$$\hat{a}_x = \frac{1}{t_n - t_0 + 1} \sum_{t=t_0}^{t_n} \ln(m_{x,t}), \quad x = 0, 1, \dots, 100+. \quad (3.56)$$

Hlavní myšlenkou vzorce (3.56) je výpočet průměru ze zlogaritmovaných měr úmrtnosti podle věku za dané časové období. Hodnoty těchto průměrů se odečtou od logaritmů měr úmrtnosti v každém roce a tím se upraví model (3.53) podle D'Amata a Russolilla (2010, s. 114)

$$\tilde{\mathbf{M}} = \ln(m_{x,t}) - \hat{a}_x = b_x k_t + \varepsilon_{x,t}, \quad x = 0, 1, \dots, 100+. \quad (3.57)$$

Matice  $\tilde{\mathbf{M}}$  (3.57) reprezentuje odchylky zlogaritmovaných měr úmrtnosti od jejich průměru. Na pravé straně ale stále zůstává neodhadnutý parametr  $b_x$ , neznámý index  $k_t$  a chybová složka  $\varepsilon_{x,t}$ . Lee a Carter (1992, s. 661) navrhnou výpočet těchto proměnných pomocí metody singulárního rozkladu.

Tuto metodu známou pod zkratkou SVD (Singular value decomposition) popisuje Watkins (2002, s. 262)

$$\tilde{\mathbf{M}} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T. \quad (3.58)$$

Matice  $\tilde{\mathbf{M}}$  typu  $m \times n$  lze rozložit podle vzorce (3.58) na ortogonální matici<sup>2</sup>  $\mathbf{U}$  typu  $n \times n$ , diagonální matici  $\mathbf{\Sigma}$  typu  $n \times m$  a ortogonální matici  $\mathbf{V}$  typu  $m \times m$ . Diagonální matici  $\mathbf{\Sigma}$  lze rozepsat (Watkins, 2002, s. 262)

---

<sup>2</sup> Ortogonální matice  $\mathbf{U}$  je čtvercová matice charakteristická tím, že její transpozice  $\mathbf{U}^T$  se rovná inverzi ortogonální matice  $\mathbf{U}^{-1}$  (Joshi, 1995, s. 66).



$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \sigma_2 & & & & & \vdots \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & \sigma_r & & \vdots \\ \vdots & & & & & 0 & \vdots \\ \vdots & & & & & & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 0 \end{pmatrix}, \quad \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0. \quad (3.59)$$

Watkins (2002, s. 262) dále uvádí, že prvky diagonální matice  $\Sigma$  (3.59) se nazývají singulární hodnoty a jsou určeny jednoznačně. Sloupce ortogonální matice  $\mathbf{U}$  z rovnice (3.58) se označují jako levostranné singulární vektory a řádky transponované ortogonální matice  $\mathbf{V}^T$  se nazývají pravostranné singulární vektory.

V praxi se často počítá zkrácená SVD (Pentreath, 2015, s. 223)

$$\tilde{\mathbf{M}} \sim \mathbf{U}_k \Sigma_k \mathbf{V}_k^T. \quad (3.60)$$

Matice  $\tilde{\mathbf{M}}$  typu  $m \times n$  se podle vzorce (3.60) přibližně rovná upravenému rozkladu SVD, kde matice  $\mathbf{U}_k$  je typu  $n \times k$ , diagonální matice  $\Sigma_k$  je typu  $k \times k$  a matice  $\mathbf{V}_k$  představuje typ  $m \times k$ . Smyslem této úpravy je, že část singulárních hodnot vysvětluje model přibližně stejně dobře, jako kdyby byly zahrnuty všechny tyto hodnoty.

**Obr. 8:** Úprava metody singulárního rozkladu při snížení dimenze diagonální matice  $\Sigma$

**Zdroj:** Pentreath (2015, s. 223) a vlastní zpracování.

Obrázek 8 graficky znázorňuje úpravu SVD při snížení počtu singulárních hodnot. Index  $k$  představuje klíčový řádek a klíčový sloupec, ve kterých dochází ke snižování dimenze diagonální matice  $\Sigma$ . Modrá plocha zobrazuje upravenou SVD.

Problematikou výpočtu SVD z hlediska přesného matematického postupu se podrobně věnují Chui a Jiang (2013, s. 116–132) nebo Howard (2010, s. 508–517). Nejjednodušeji se SVD vypočítá v programu Matlab (2012b) příkazem

$$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \text{svd}(\mathbf{M}). \quad (3.61)$$

Jelikož programy všeobecně neumožňují užití speciálních znaků je v předpisu (3.61)  $\tilde{\mathbf{M}}$  nahrazeno  $\mathbf{M}$  a  $\Sigma$  nahrazena  $\mathbf{S}$ . Jedná se pouze o formální úpravu, která nemá na výpočet žádný vliv. Jednotlivé prvky SVD již byly popsány v rovnici (3.58).

V této fázi je možné odhadnout parametr  $b_x$  a neznámý index  $k_t$  Lee-Carterova modelu ze vztahu (Koissi a kol., 2005, s. 4)

$$\hat{b}_x = u_{xk}, \quad k=1, \quad x=0,1,\dots,100+, \quad (3.62)$$

$$\hat{k}_t = \sigma_1 v_{tk}, \quad k=1. \quad (3.63)$$

K výpočtu je zapotřebí rozklad zkrácené SVD podle vztahu (3.60). Smysl této úpravy spočívá ve snížení dimenze diagonální matice  $\Sigma$ . Počet singulárních hodnot je snížen na jednu a to na tu, která je nejvyšší. Ze vztahu (3.59) platí, že touto nejvyšší hodnotou je právě první z nich. Tím je jednoznačně určen první řádek i sloupec jako klíčový. Z obrázku 8 je poté již jednoduché odvodit, že ortogonální matice  $U$  a  $V$  jsou redukovány na jeden sloupec transformované matice  $U_1$  a  $V_1$ . Jelikož se jedná o jeden řádek matice, je přesnější místo pojmu matice používat označení vektor  $u_1$  a  $v_1$ . Ze vztahu (3.62) lze odhadnout parametr věkově specifických změn v úmrtnosti, který je v principu roven vektoru  $u_1$ . Podle vzorce (3.63) se vypočítá index úrovně úmrtnosti jako součin první singulární hodnoty a vektoru  $v_1$ . Jako poslední zbývá u Lee-Carterova modelu odhadnout chybovou složku, kterou lze získat úpravou rovnice (3.53) podle D'Amata a Russolilla (2010, s. 114)

$$\hat{\varepsilon}_{x,t} = \ln(m_{x,t}) - (\hat{a}_x + \hat{b}_x \hat{k}_t), \quad x=0,1,\dots,100+. \quad (3.64)$$

Odhad chybové složky (3.64) se vypočítá jako rozdíl skutečných zlogaritmovaných měr úmrtnosti a odhadnutých zlogaritmovaných měr úmrtnosti. Tyto odhadnuté zlogaritmované míry úmrtnosti jsou vyjádřeny jako odhady průměrů zlogaritmovaných měr úmrtnosti podle věku a k nim jsou přičteny odhady věkově specifických změn v úmrtnosti vynásobené s odhadnutým indexem úrovně úmrtnosti.

Lee a Cater (1992, s. 661) tvrdí, že  $a_x$  a  $b_x$  jsou věkově specifické konstanty nezávislé na čase. Naopak závislý na čase je index  $k_t$ , který tvoří v modelu proměnlivou složku. Tento index jako jediný slouží pro modelování budoucího vývoje úmrtnosti. Lee a Cater (1992, s. 663) doporučují pro modelování směřování indexu  $k_t$  tzv. Box-Jenkinsovu metodologii. Zmiňují se především o modelech ARIMA. Tyto modely slouží převážně pro modelování dlouhých časových řad a jejich odhad má stochastický charakter. Podrobněji o těchto modelech v podkapitole 4.2.

### 3.3 Vybrané modely při studiu očekávaného vývoje plodnosti

Stejně jako úmrtnost patří i plodnost mezi základní komponenty kohortně komponentní metody bez migrace. Počátek studia úmrtnosti spadá do 17. století a je spojeno se jménem Johna Graunta. Plodnost byla analyzována mnohem později. Studium plodnosti podle věku je analyzováno až na konci 19. století a dává se do souvislosti především se jménem Richarda Böckha. Dále se plodností zabývali Robert René Kuczynski, Alfred James Lotka, Karl Pearson, Ansley Johnson Coale a další (Pavlík a kol., 1986, s. 171, 286).

Podkladem pro analýzu plodnosti jsou míry plodnosti podle věku. Při studiu historických údajů byla data za plodnost často nestabilní a nespolehlivá. Docházelo ke snahám nalézt zákonitosti, které by dokázaly popsat plodnost. Vyvinula se celá řada nástrojů pro vyrovnání plodnosti, které nepostrádají smysl ani v dnešní době. Používají se především pro modelování plodnosti malých územních celků. Pro účely této práce budou použity analytické nástroje

vyrovnání křivek plodnosti. Mezi tyto nástroje patří například Beta funkce. Další analytické nástroje vyrovnání plodnosti popisují například Peristera a Kostaki (2007, s. 145–149).

Beta funkce se používá převážně v matematice pro výpočet některých složitějších integrálů. Dále se aplikuje ve statistice ve formě beta rozdělení. V demografii ji použil Roland Pressat (1995, s. 62–67) pro rozložení měr plodnosti podle věku. Všiml si, že rozložení těchto měr má tvar blízký Beta funkci. Hlavní myšlenkou je, že Beta funkce má pravidelný tvar, ale rozložení měr plodnosti podle věku může být vychýlené a prostřednictvím Beta funkce se tyto míry vyrovnají. Jak již bylo řečeno u úmrtnosti, bude i pro plodnost uvažován jednoletý věkový interval a na základě toho budou i všechny následující vzorce upraveny pro jednoletý věkový interval.

Klasický výpočet Beta funkce (Chaudhry a kol., 1997, s. 20)

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt, \quad x > 0, \quad y > 0. \quad (3.65)$$

V rovnici (3.65) představuje  $B(x, y)$  Beta funkci, kde  $x$  a  $y$  jsou její parametry tvaru, dále je definované  $t$  jako spojitá proměnná. Beta funkce je určitý integrál definovaný od nuly do jedné a specifikovaný pro první kvadrant, tzn. vyjádřený pro  $x$  a  $y$ , které jsou větší než nula. Beta funkci lze převést na gama funkci. Hodnoty gama funkce, které nabývají přirozených čísel, lze již vyčíslit. Beta funkce se převede na gama funkci podle vztahu (Chaudhry a kol., 1997, s. 20)

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \quad (3.66)$$

Ve vzorci (3.66) symbolizuje  $\Gamma(x)$ ,  $\Gamma(y)$  a  $\Gamma(x+y)$  gama funkci parametrů  $x$  a  $y$ . Transformací Beta funkce na gama funkci se zjednodušuje výpočet. Podrobněji se celým matematickým postupem výpočtu gama funkce zabývá Hassani (2009, s. 318–319).

Aby bylo možné vyrovnat míry plodnosti podle věku, je potřeba rovnici (3.65) upravit (Pressat, 1995, s. 63)

$$f(x) = D(\beta) \frac{1}{B(a, b)} \frac{(x-\alpha)^{a-1} (\beta-x)^{b-1}}{(\beta-\alpha)^{a+b-1}}, \quad 1 < a < b. \quad (3.67)$$

V rovnici (3.67) představuje  $x$  věk,  $D(\beta)$  úhrnnou plodnost,  $\alpha$  minimální přesný věk plodnosti,  $\beta$  maximální přesný věk plodnosti,  $B(a, b)$  symbolizuje Beta funkci, viz transformace (3.66), kde parametry  $x$  a  $y$  jsou nahrazeny parametry  $a$  a  $b$ . Jedná se pouze o formální úpravu. Oba tyto parametry budou definovány později. Parametry  $a$  a  $\beta$  jsou určeny z rovnice (3.67) jednoznačně (Pressat, 1995, s. 65)

$$\alpha = 15, \quad (3.68)$$

$$\beta = 50. \quad (3.69)$$

Ve vzorci (3.68) symbolizuje  $\alpha$  minimální přesný věk plodnosti, který je stanoven na 15 let a ve vzorci (3.69) značí  $\beta$  maximální přesný věk plodnosti, který je určen na 50 let. Úhrnná plodnost se vypočítá (Pressat, 2008, s. 95; Pressat, 1995, s. 64)

$$D(\beta) = \sum_{15}^{49} f_x. \quad (3.70)$$

Vzorec (3.70) představuje součet měr plodnosti  $f_x$  v dokončeném věku 15 až 49 let. Jedná se o věk, ve kterém ženy realizují svou plodnost.

Pro následující text je zapotřebí definovat pojem náhodná veličina. Tento pojem je podrobně popsán v podkapitole 4.2. Střední hodnota a rozptyl náhodné veličiny  $X$  se získá (Pressat, 1995, s. 65)

$$m(X) = m\left(\frac{x - \alpha}{\beta - \alpha}\right) = \frac{m(x) - 15}{35}, \quad (3.71)$$

$$s^2(X) = s^2\left(\frac{x - \alpha}{\beta - \alpha}\right) = \frac{s^2(x)}{35^2}. \quad (3.72)$$

Ve vzorci (3.71) značí obecně  $m$  střední hodnotu a  $m(x)$  představuje průměrný věk matek. Ve formulaci (3.72) symbolizuje obecně  $s^2$  rozptyl a  $s^2(x)$  označuje rozptyl průměrného věku matek. Oba vzorce mají společné symboly věk  $x$ ,  $\alpha$  minimální přesný věk plodnosti a  $\beta$  maximální přesný věk plodnosti. Střední hodnota náhodné veličiny  $X$  (3.71) se vypočítá jako průměrný věk matek snížený o 15 let, tzn. převede se na průměrný počet let, za které se patnáctiletá žena stane matkou a vydělí se šířkou variačního rozpětí maximálního a minimálního přesného věku plodnosti tj. 35 let. Rozptyl náhodné veličiny  $X$  (3.72) se rovná rozptylu průměrného věku matek vydělený druhou mocninou variačního rozpětí věku. Aby bylo možné vypočítat vzorce (3.71) a (3.72) je nejprve nutné vypočítat průměrný věk matek a rozptyl průměrného věku matek (Pavlík a kol., 1986, s. 363; Mathcentre, 2013, s. 1–2)

$$m(x) = \frac{\sum_{15}^{49} (x + 0,5) f_x}{\sum_{15}^{49} f_x}, \quad (3.73)$$

$$s^2(x) = \frac{\sum_{15}^{49} (x + 0,5)^2 f_x}{\sum_{15}^{49} f_x} - \left( \frac{\sum_{15}^{49} (x + 0,5) f_x}{\sum_{15}^{49} f_x} \right)^2, \quad (3.74)$$

kde  $x$  je věk a  $f_x$  jsou míry plodnosti podle věku. Průměrný věk matek se vypočítá z rovnice (3.73), kde váhy představují míry plodnosti podle věku. V čitateli je součet vah vynásobený středem jednoletého věkového intervalu a ve jmenovateli je úhrnná plodnost. Rozptyl průměrného věku matky udává vzorec (3.74). Vychází z klasického výpočetního vzorce rozptylu, kdy první člen zahrnuje v čitateli součet druhých mocnin středu věkového intervalu vynásobeného s vahami a ve jmenovateli je součet těchto vah. Druhý člen představuje druhou mocninu průměrného věku matek, viz vzorec (3.73).

Nyní jsou již k dispozici všechny potřebné podklady pro výpočet parametrů  $a$  a  $b$  (Pressat, 1995, s. 64)

$$a = \frac{[1 - m(X)] m^2(X)}{s^2(X)} - m(X), \quad (3.75)$$

$$b = \frac{[1 - m(X)]^2 m(X)}{s^2(X)} - [1 - m(X)]. \quad (3.76)$$

Parametr  $a$  (3.75) a  $b$  (3.76) jsou dva ukazatele tvaru rozložení intenzit plodnosti podle věku. Do jejich výpočtu vstupuje střední hodnota náhodné veličiny  $X$  (3.71) a rozptyl náhodné veličiny  $X$  (3.72).

Pokud by se odhadl vývoj úhrnné plodnosti, budoucí směřování průměrného věku matek a vývoj rozptylu průměrného věku matek, pak by bylo možné odhadnout celou křivku plodnosti pomocí Beta funkce. Projekce parametrů funkcí má nespornou výhodu především proto, že výsledné rozložení měr plodnosti podle věku je již vyrovnané. Odhad těchto parametrů je možný buď expertním odhadem nebo na základě analýzy časových řad. Tato problematika bude popsána v kapitole 4.

Dalším modelem, který lze užít pro dílčí prognózu plodnosti, je Lee-Carterův model plodnosti. V roce 1993 upravil Ronald Demos Lee (1993, s. 187–202) Lee-Carterův model úmrtnosti a navrhl jeho zpracování pro plodnost. Jedná se o extrapolační metodu, která využívá analýzy časových řad. On sám ji použil pro prognózu plodnosti ve Spojených státech amerických. Celý Lee-Carterův model plodnosti lze zapsat (Lee, 1993, s. 190)

$$f_{x,t} = a_x + b_x f_t + \varepsilon_{x,t}, \quad x = 15, \dots, 49. \quad (3.77)$$

Ve vzorci (3.77) představují  $f_{x,t}$  míry plodnosti podle věku  $x$  v roce  $t$ ,  $a_x$  je průměrný tvar plodnosti podle věku  $x$ ,  $b_x$  označuje věkově specifické změny v plodnosti podle věku  $x$ ,  $f_t$  určuje index úrovně plodnosti v roce  $t$  a  $\varepsilon_{x,t}$  je chybová složka ve věku  $x$  v roce  $t$ , u které se předpokládá normální rozdělení. K získání jednoznačného řešení Lee-Carterova modelu uvádí Lee (1993, s. 190) podmínky

$$\sum_{x=15}^{49} b_x = 1, \quad (3.78)$$

$$\sum_{t=t_0}^{t_n} f_t = 0. \quad (3.79)$$

Omezení (3.78) značí, že součet věkově specifických změn v plodnosti se rovná jedné. Druhé omezení (3.79) symbolizuje nulový součet indexů plodnosti. Součet indexů je stanoven od počátku časového období  $t_0$  po jeho konec  $t_n$ .

První fází výpočtu Lee-Carterova modelu je odhadnout průměrný tvar plodnosti podle věku (Lee, 1993, s. 190)

$$\hat{a}_x = \frac{1}{t_n - t_0 + 1} \sum_{t=t_0}^{t_n} f_{x,t}, \quad x = 15, \dots, 49. \quad (3.80)$$

Principem výpočtu odhadu průměrného tvaru plodnosti (3.80) je klasický průměr měr plodnosti ve věku  $x$  za dané časové období od  $t_0$  až do  $t_n$ . Tímto způsobem spočítaný odhad  $a_x$  se odečte od měr plodnosti v každém roce (Myrskylä a kol., 2013, s. 50)

$$\mathbf{F} = f_{x,t} - \hat{a}_x = b_x f_t + \varepsilon_{x,t}, \quad x = 15, \dots, 49. \quad (3.81)$$

Matice  $\mathbf{F}$  (3.81) představuje odchylky měr plodnosti od průměru těchto měr. Na levé straně této rovnice zůstává neodhadnutý parametr  $b_x$  a neznámý index  $f_t$ . Lee (1993, s. 190) počítá tyto proměnné prostřednictvím singulárního rozkladu. Princip tohoto odhadu byl popsán v podkapitole 3.2. Jediný rozdíl spočívá v tom, že se místo transformované matice  $\tilde{\mathbf{M}}$  nahradí maticí  $\mathbf{F}$ . Poslední zbývá odhadnout z rovnice (3.81) chybovou složku, její výpočet je inspirovaný úpravou rovnice úmrtnosti od D'Amata a Russolilla (2010, s. 114)

$$\hat{\varepsilon}_{x,t} = f_{x,t} - (\hat{a}_x + \hat{b}_x \hat{f}_t), \quad x = 15, \dots, 49. \quad (3.82)$$

Odhad chybové složky (3.82) podle věku  $x$  v roce  $t$  je vymezen rozdílem skutečných a odhadnutých měr plodnosti podle věku  $x$  v roce  $t$ .

Věkově specifické parametry jsou nezávislé na čase, pouze index  $f_t$  na čase závisí. Tento index je potřeba odhadnout do budoucna a stejně jako u Lee-Carterova modelu úmrtnosti je i u analýzy plodnosti užita Box-Jenkinsova metodologie. Tyto modely budou podrobněji probrány v podkapitole 4.2. Lee (1993, s. 190) navrhl také logaritmickou verzi Lee-Carterova modelu plodnosti

$$\ln(f_{x,t}) = a_x + b_x f_t + \varepsilon_{x,t}, \quad x = 15, \dots, 49. \quad (3.83)$$

Levá strana rovnice (3.83) představuje logaritmus měr plodnosti podle věku  $x$  v roce  $t$ . Pravá strana rovnice obsahuje průměrný tvar zlogaritmovaných měr plodnosti podle věku  $x$  pod zkratkou  $a_x$ ,  $b_x$  symbolizuje věkově specifické změny v plodnosti podle věku  $x$ , index úrovně plodnosti v roce  $t$  je označen  $f_t$  a  $\varepsilon_{x,t}$  reprezentuje chybovou složku podle věku  $x$  v roce  $t$ . Odhad všech parametrů včetně podmínek je v principu totožný s Lee-Carterovým modelem úmrtnosti, který byl již popsán v podkapitole 3.2.

### 3.4 Přístup k analýze migrace

Doposud se mluvilo o modelech úmrtnosti a plodnosti jako o podkladu pro kohortně komponentní metodu bez migrace. Otázkou je, jakým způsobem analyzovat migraci, která patří k nejproblematictějším složkám kohortně komponentní metody především kvůli nedostatku dat a rovněž proto, že její vývoj je navázán na aktuální socio-ekonomické a politické klima. Na úrovni České republiky existují data za přistěhovalé a vystěhovalé podle věku a pohlaví. Na úrovni krajů uveřejňuje Český statistický úřad pouze celkové počty přistěhovalých a vystěhovalých podle pohlaví. Problémem je, že tyto počty nejsou členěny dle věku.

Klosterman (1990, s. 91) v této souvislosti zavádí nový pojem tzv. **reziduální migrační saldo**, které je charakteristické tím, že se získá výpočtem z bilanční rovnice. V principu se jedná o odhad migračního salda vyjádřeného z bilanční rovnice (3.32)

$$\widehat{NM} = (P_t - P_0) - (B - D). \quad (3.84)$$

Odhad reziduálního migračního salda pod označením  $\widehat{NM}$  se vypočítá z rovnice (3.84) jako rozdíl celkové změny počtu žijících obyvatel mezi výchozím okamžikem a rokem  $t$  a celkového přirozeného přírůstku od výchozího okamžiku až do roku  $t$  (Klosterman, 1990, s. 91).

Hinde (1998, s. 194) celý postup výpočtu reziduálního migračního salda označuje jako **nepřímá metoda výpočtu migračního salda** a dále tento postup rozšiřuje o odhad věkově specifického migračního salda. Nepřímá metoda výpočtu migračního salda úzce souvisí s kohortně komponentní metodou bez migrace, kdy se struktura počtu žijících v roce  $t$  posouvá po generacích do roku  $t + I$  za působení úmrtnosti zahrnuté v pravděpodobnostech přežití. Rozdíl ale spočívá v tom, že nepřímá metoda výpočtu migračního salda analyzuje skutečnou časovou řadu za minulá období a jejím výstupem je odhad reziduálního migračního salda za toto období, naopak kohortně komponentní metoda transformuje aktuální věkovou strukturu do budoucna a jejím výstupem jsou odhadované počty žijících.

Techniku nepřímé metody výpočtu migračního salda je možné rozdělit do tří kroků. Prvním krokem je odhadnout reziduální migrační saldo pro generace  $z - 1$  až  $z - 99$ . Druhý krok spočívá v odhadu reziduálního migračního salda pro poslední otevřený interval. Posledním krokem výpočtu je odhadnout migrační saldo z generace narozených tedy  $z$ . Jednotlivé kroky budou nyní detailněji rozebrány a budou aplikovány na transversální úmrtnostní tabulky. Pro snadnější orientaci ve vzorcích je v tabulce 3 uveden seznam symbolů.

**Tab. 3: Seznam symbolů použitých u nepřímé metody výpočtu migračního salda**

Symbol	Popis
$f$	ženy
$m$	muži
$x$	věk
$t$	rok
$z$	generace
$k$	krok transformace
${}^{z-1-x}_t \widehat{NM}^f$	odhad reziduálního migračního salda žen z generace $z - 1 - x$ v roce $t$ spočítaného z druhého hlavního souboru
${}^{z-1-x}_t S_x^f$	pravděpodobnost přežití žen v dokončeném věku $x$ od počátečního stavu v roce $t$ ke koncovému stavu v roce $t$ z generace $z - 1 - x$
${}^{z-1-x}_t L_x^f$	tabulkový počet žijících žen v dokončeném věku $x$ v roce $t$ z generace $z - 1 - x$
${}^{z-1-x}_t P_x^f$	počet žijících žen v dokončeném věku $x$ k výchozímu okamžiku v roce $t$ z generace $z - 1 - x$
${}^{t+\frac{1}{2}}_t f_x$	míry plodnosti podle věku $x$ ke střednímu stavu roku $t$ spočítané z třetího hlavního souboru
${}^{z-1-x}_{t+\frac{1}{2}} f$	míry plodnosti z generace žen $z - 1 - x$ ke střednímu stavu roku $t$ spočítané z druhého hlavního souboru
$\phi$	podíl děvčat při narození
${}_t^z l_0^f$	kořen tabulky žen v roce $t$ z generace $z$

**Zdroj:** Kučera (1998, s. 75–78), Hoque a kol. (2013, s. 12), vlastní zpracování.

**První krok** výpočtu spočívá v odhadu reziduálního migračního salda ve druhém hlavním souboru pro generace  $z - 1$  až  $z - 99$  (Hinde, 1998, s. 195)

$${}^{z-1-x}_t \widehat{NM}^f = {}^{z-1-x}_{t+1} P_{x+1}^f - {}^{z-1-x}_t P_x^f * {}^{z-1-x}_t S_x^f, \quad x = 0, 1, \dots, 98, \quad (3.85)$$

$${}^{z-1-x}_t \widehat{NM}^m = {}^{z-1-x}_{t+1} P_{x+1}^m - {}^{z-1-x}_t P_x^m * {}^{z-1-x}_t S_x^m, \quad x = 0, 1, \dots, 98. \quad (3.86)$$

Hlavní myšlenkou při stanovení odhadu reziduálního migračního salda je rozdíl mezi skutečným a odhadnutým počtem žijících v roce  $t + 1$  zvlášť pro ženy (3.85) a muže (3.86). Odhadnutý počet žijících v roce  $t + 1$  se získá vynásobením pravděpodobností přežití s počtem žijících v roce  $t$ . Pravděpodobnosti přežití jsou zpracovány ze skutečných úmrtnostních tabulek v daném roce a jejich výpočet byl již definován v podkapitole 3.1.1. Podle zvoleného přístupu jsou vybrány vzorce (3.2) a (3.3) nebo (3.4) a (3.5).

**Druhým krokem** výpočtu je odhad reziduálního migračního salda v posledním otevřeném intervalu v rámci druhého hlavního souboru, který lze odvodit z obrázku 7 kategorie  $b$  a je inspirován vzorcem Klostermana (1990, s. 76)

$${}^{z-1-(x+)}_t \widehat{NM}^f = {}^{z-1-(x+)}_{t+1} P_{x+}^f - {}^{z-1-(x+)}_t P_{x+}^f * {}^{z-1-(x+)}_t S_{x+}^f - {}^{z-x}_t P_{x-1}^f * {}^{z-x}_t S_{x-1}^f, \quad x = 100, \quad (3.87)$$

$${}^{z-1-(x+)}_t \widehat{NM}^m = {}^{z-1-(x+)}_{t+1} P_{x+}^m - {}^{z-1-(x+)}_t P_{x+}^m * {}^{z-1-(x+)}_t S_{x+}^m - {}^{z-x}_t P_{x-1}^m * {}^{z-x}_t S_{x-1}^m, \quad x = 100. \quad (3.88)$$

Poslední otevřený interval je charakteristický tím, že zahrnuje více generací. Princip zpracování zůstává stejný jako v prvním kroku výpočtu. Jedná se o rozdíl skutečného počtu žijících a odhadnutého počtu žijících k výchozímu okamžiku v roce  $t + 1$ . Odhad počtu žijících je ale trochu složitější, protože do posledního otevřeného intervalu vstupuje generace  $z - 100$  a generace přeživších z posledního otevřeného intervalu. Pravděpodobnosti přežití v nejvyšších věcích se vypočítají ze vzorců (3.10) až (3.13) za použití skutečných úmrtnostních tabulek. Odhad počtu žijících ve věku 100 a více let v roce  $t + 1$  se vypočítá jako součet počtu žijících v roce  $t$  ve věku 100 a více let vynásobený s pravděpodobnostmi přežití v tomto věku a počtu žijících v roce  $t$  ve věku 99 vynásobený s pravděpodobnostmi přežití ve věku 99. Rovnice (3.87) a (3.88) se liší o pohlaví.

**Třetím a posledním krokem** výpočtu je odhadnout reziduální migrační salda pro generaci narozených. Jestliže jsou k dispozici počty narozených podle pohlaví, je situace nejjednodušší a přechází se rovnou na výpočet rovnic (3.93) a (3.94). Pokud jsou k dispozici míry plodnosti v rámci druhého hlavního souboru, poté se počet narozených vypočítá (Kučera, 1998, s. 76; Roubíček, 1996, s. 212)

$${}_t B^v = \sum_{x=15}^{49} \frac{({}^{z-1-x}P_x^f + {}^{z-1-x}P_{x+1}^f)}{2} * {}^{z-1-x}f_{t+\frac{1}{2}}. \quad (3.89)$$

Podstatou vzorce (3.89) je vynásobení středního stavu počtu žijících s měrami plodnosti ke středu intervalu roku  $t$  z generace žen  $z - 1 - x$ . Tímto způsobem se vypočítají počty narozených z generace žen  $z - 1 - x$  v roce  $t$  a jejich součtem se získá celkový počet narozených. Při znalosti měr plodnosti zahrnutých ve třetím hlavním souboru se počty narozených vyjádří (Kučera, 1998, s. 76; Roubíček, 1996, s. 212)

$${}_t B^v = \sum_{x=15}^{49} \frac{({}^{z-1-x}P_x^f + {}^{z-x}P_x^f)}{2} * {}^{t+\frac{1}{2}}f_x. \quad (3.90)$$

Rozdíl mezi oběma vzorci spočívá ve výpočtu středního stavu, zatímco ve vyjádření (3.89) se zachovávají generace, ve vzorci (3.90) je shodný věk počtu žijících. Následně se počty narozených rozdělí dle pohlaví (Kučera, 1998, s. 77)

$${}_t \hat{B}^{v,f} = \phi * {}_t B^v, \quad (3.91)$$

$${}_t \hat{B}^{v,m} = (1 - \phi) * {}_t B^v. \quad (3.92)$$

Odhad počtu živě narozených dívek (3.91) a chlapců (3.92) se vypočítá podle stejného postupu jako ve vzorcích (3.24) a (3.25).

Odhad migračního salda od narození do výchozího okamžiku roku  $t + 1$  se vypočítá (Hinde, 1998, s. 196)

$$\widehat{{}_t NM_0^f} = {}^{z-1}P_0^f - {}_t \hat{B}^{v,f} * \frac{{}^{z-1}L_0^f}{{}_t L_0^f}, \quad (3.93)$$

$$\widehat{{}_t NM_0^m} = {}^{z-1}P_0^m - {}_t \hat{B}^{v,m} * \frac{{}^{z-1}L_0^m}{{}_t L_0^m}. \quad (3.94)$$

Odhad reziduálního migračního salda od přesného věku nula do dokončeného věku nula se stanoví zvlášť pro ženy (3.93) a zvlášť pro muže (3.94). Princip výpočtu spočívá v rozdílu skutečného a odhadnutého počtu žijících v dokončeném věku nula. Odhad počtu žijících



v dokončeném věku nula se vypočítá vynásobením odhadu počtu živě narozených s podílem tabulkového počtu žijících žen v dokončeném věku nula a tabulkovým počtem dožívajících se přesného věku nula.

Analýza migrace bude v této práci založena na odhadnutém reziduálním migračním saldu podle věku a pohlaví. Jelikož se jedná o odhady migračního salda a ne o skutečné počty, jsou odhady zatíženy chybou. Kvůli této skutečnosti nebudou použity metody na vyrovnání migračních křivek<sup>3</sup>, jejichž parametry by bylo možné na základě minulého průběhu časové řady projektovat do budoucna. Ve výsledku by tak docházelo ke kumulaci chyb, protože odhad parametrů by znamenal další rozšíření chyb.

Časová řada věkově specifického migračního salda za minulá období může být analyzována nejjednodušeji pomocí střední hodnoty, kdy v jednotlivých věcích za minulá období se vypočítá průměr nebo medián. Tím se získá nejpravděpodobnější věková struktura migračního salda. Johnson a Winkler (2015, s. 1068–1070) užívají k analýze struktury migračního salda **medián**. Výhoda užití mediánu oproti průměru spočívá v tom, že medián je méně citlivý na odlehlá pozorování. Otázkou je, jakým způsobem projektovat věkovou strukturu do budoucna. Štyglarová a kol. (2010, s. 3) zahrnují do svých prognóz konstantní migrační saldo. Stejně tak lze jako konstantu uvažovat odhadnutou věkovou strukturu pomocí mediánu a následně ji celou projektovat do budoucna. Nevýhodou tohoto postupu je, že není uvažován žádný trend, který by vycházel z průběhu časové řady za minulá období. Tento problém řeší například Lee-Carterův model čisté migrace.

**Lee-Carterův model čisté migrace** lze zapsat (García-Guerrero, 2016, s. 119; Lee, 1993, s. 192)

$$g_{x,t} = a_x + b_x g_t + \varepsilon_{x,t}, \quad x = 0, 1, \dots, 100+. \quad (3.95)$$

V rovnici (3.95) představuje  $g_{x,t}$  migrační saldo podle věku  $x$  v roce  $t$ ,  $a_x$  symbolizuje průměrný tvar křivky migračního salda,  $b_x$  charakterizuje věkově specifické změny migračního salda podle věku  $x$ ,  $g_t$  specifikuje index úrovně migračního salda v roce  $t$  a  $\varepsilon_{x,t}$  je chybová složka ve věku  $x$  v roce  $t$ . Výpočet Lee-Carterova modelu čisté migrace je shodný s Lee-Carterovým modelem plodnosti viz kapitola 3.3.

Rovněž je možné prostřednictvím Lee-Carterova modelu prognózovat zvlášť vystěhované a zvlášť přistěhované. Počty přistěhovaných nejsou závislé na věkové struktuře, a proto není vhodné používat intenzity tohoto ukazatele (Kučera, 1998, s. 81). **Lee-Carterův model emigrace** se vypočítá (Raymer a kol., 2012, s. 1; Lee a Carter, 1992, s. 661)

$$\ln(e_{x,t}) = a_x + b_x e_t + \varepsilon_{x,t}, \quad x = 0, 1, \dots, 100+. \quad (3.96)$$

V rovnici (3.96) označuje  $e_{x,t}$  míry vystěhování podle věku  $x$  v roce  $t$ ,  $a_x$  představuje průměrný tvar zlogaritmovaných měr vystěhování podle věku  $x$ ,  $b_x$  symbolizuje věkově specifické změny měr vystěhování podle věku  $x$ ,  $e_t$  charakterizuje index úrovně měr vystěhování v roce  $t$  a  $\varepsilon_{x,t}$  je chybová složka ve věku  $x$  v roce  $t$ .

<sup>3</sup> Modelováním křivek migrace se zabývali například Rogers a Castro (1981, s. 6–7), kteří vytvořili jedenáctiparametrový model migrace, kdy věkový profil je definován sedmi parametry a úroveň zbývajících čtyřmi parametry.

**Lee-Carterův model imigrace** se vypočítá (Raymer a kol., 2012, s. 1; Lee a Carter, 1992, s. 661)

$$\ln(I_{x,t}) = a_x + b_x I_t + \varepsilon_{x,t}, \quad x = 0, 1, \dots, 100+. \quad (3.97)$$

V rovnici (3.97) symbolizuje  $I_{x,t}$  počty přistěhovalých podle věku  $x$  v roce  $t$ ,  $a_x$  značí zlogaritmovaný průměrný tvar křivky počtu přistěhovalých podle věku  $x$ ,  $b_x$  charakterizuje věkově specifické změny počtu přistěhovalých podle věku  $x$ ,  $I_t$  označuje index úrovně počtu přistěhovalých v roce  $t$  a  $\varepsilon_{x,t}$  je chybová složka ve věku  $x$  v roce  $t$ . Výpočet Lee-Carterova modelu emigrace a imigrace je shodný s Lee-Carterovým modelem úmrtnosti viz kapitola 3.2.

## Kapitola 4

### Vybrané přístupy k analýze časových řad s potenciálním využitím pro demografické prognózování

Analýza časových řad se na základě minulého vývoje snaží predikovat budoucnost. Rozlišují se krátké a dlouhé časové řady. **Dlouhé časové řady** „(...) jsou sledovány v roční frekvenci (...)“ naopak **krátké časové řady** „(...) jsou sledovány ve frekvencích kratších než jeden rok (...)“ (Arlt a Arltová, 2003, s. 12). Časovou řadu lze rozložit na čtyři složky: trendovou, cyklickou, sezónní a reziduální. **Trendová složka** ( $T_t$ ) vyjadřuje dlouhodobé tendence. **Sezónní složka** ( $S_t$ ) představuje pravidelné kolísání během jednoho roku, které se každoročně opakuje. Příkladem jsou sezónní práce, změny ve sňatečnosti, kdy kvůli pověře jsou sníženy počty sňatků v květnu atd. **Cyklická složka** ( $C_t$ ) zachycuje stejně jako sezónní složka pravidelné kolísání. Rozdíl ale spočívá v délce jednoho cyklu, který je delší než jeden rok. Příkladem mohou být hospodářské cykly, změny klimatu a další. Trendová, sezónní a cyklická složka patří mezi systematické složky. Poslední složkou při dekompozici je tzv. **reziduální složka** ( $E_t$ ). Jedná se o nesystematickou složku, která tvoří náhodné výkyvy. Z časové řady se získá, když se odstraní trendová složka, sezónní složka a cyklická složka (Cipra, 1986, s. 16–17).

Cipra (1986, s. 18) a Arlt a kol. (2002, s. 20) představují dva typy rozkladů časové řady tzv. aditivní dekompozice a multiplikativní dekompozice. **Aditivní dekompozice** je charakteristická součtem všech čtyř složek časové řady

$$y_t = T_t + C_t + S_t + E_t, \quad (4.1)$$

V rovnici (4.1) symbolizuje  $y_t$  skutečnou časovou řadu, která se skládá ze složek vyjádřených v absolutních hodnotách. Používá se tehdy, jestliže je rozptyl v čase konstantní. **Multiplikativní dekompozice** se vypočítá vynásobením všech čtyř složek časové řady

$$y_t = T_t * C_t * S_t * E_t. \quad (4.2)$$

V rozkladu (4.2) je zpravidla pouze trendová složka vyjádřená v absolutní hodnotě. Zbylé složky jsou formulovány v relativních hodnotách a jsou vyjádřeny jako bezrozměrné. Aplikuje se tehdy, pokud se rozptyl zvětšuje nebo se liší v čase. Rozkladem časové řady s přímým využitím v demografii se zabývají například Myšáková (2011, s. 21–77) a Kačerová (2004, s. 186–198).

Jelikož v analytické části této práce jsou k dispozici pouze roční data, odpadá z rovnice sezónní složka. Cyklická složka se zjišťuje velmi problematicky, zpravidla pomocí klouzavých průměrů. Pokud je k dispozici málo údajů, je prakticky nemožné cyklickou složku identifikovat.

Z tohoto důvodu bude analýza zaměřena na trendovou funkci. Pokud je k dispozici dlouhá časová řada, je možné uplatnit složitější postupy, například Box-Jenkinsovu metodologii. Více o tomto přístupu v podkapitole 4.2.

## 4.1 Vybrané trendové funkce

Identifikace trendu tvoří základní složku při rozkladu časové řady. Cílem je zjistit dlouhodobé tendence pomocí analytických funkcí. Trendové funkce se používají především pro odhad budoucnosti, kdy za pomoci extrapolace křivky lze predikovat očekávaný vývoj časové řady. Při aditivní dekompozici a za předpokladu již odstraněné nebo neexistující sezónní a cyklické složky lze časovou řadu rozložit na funkci trendu a reziduální složky (Cipra, 1986, s. 29)

$$y_t = T_t + \varepsilon_t. \quad (4.3)$$

V časové řadě  $y_t$  (4.3) je trendová složka  $T_t$  charakterizována analytickou funkcí a tvoří systematickou část dekompozice. Druhá část je nesystematická a je symbolizována v dekompozici reziduální složkou  $\varepsilon_t$ , která by měla vykazovat vlastnosti procesu bílého šumu.

**Procesem bílého šumu** lze nazvat takový stochastický proces, který tvoří řadu nekorelovaných náhodných veličin a jeho základní charakteristiky jsou (Arlt a Arltová, 2003, s. 34; Arlt a kol., 2002, s. 21)

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad E(\varepsilon_t) = 0, \quad D(\varepsilon_t) = \sigma_\varepsilon^2, \quad C(\varepsilon_t, \varepsilon_{t-k}) = 0, \quad k \neq 0, \quad (4.4)$$

kde  $\varepsilon_t$  je reziduální složka v čase  $t$ ,  $\varepsilon_{t-k}$  znázorňuje reziduální složku zpožděnou o  $k$  pozorování,  $\sigma_\varepsilon^2$  symbolizuje rozptyl reziduální složky,  $N$  představuje normální rozdělení,  $E$  značí střední hodnotu,  $C$  vyjadřuje kovarianci. Reziduální složka by podle definice (4.4) měla splňovat určité vlastnosti, mezi které patří, že střední hodnota reziduální složky by měla mít přibližně normální rozdělení se střední hodnotou nula a konstantním rozptylem, dále by měla mít nulovou autokovarianci. Při splnění těchto vlastností, se reziduální složka nazývá bílým šumem (Cipra, 1986, s. 17).

Existuje celá řada analytických funkcí, které se užívají při konstrukci trendu  $T_t$ . K nejjednodušším z nich patří **konstantní trend** (Cipra, 1986, s. 30)

$$T_t = \beta_0. \quad (4.5)$$

Jedná se o **polynomický trend** nultého stupně. Trend v čase  $t$  je podle vztahu (4.5) závislý pouze na parametru  $\beta_0$ . V grafu by byl znázorněn přímkou, která je rovnoběžná s jeho horizontální osou. Polynomické trendy vyšších stupňů se konstruují (Medhi, 1992, s. 317)

$$T_t = \beta_0 + \beta_1 t, \quad (4.6)$$

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2, \quad (4.7)$$

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3, \quad (4.8)$$

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \dots + \beta_k t^k. \quad (4.9)$$

Polynom prvního stupně (4.6) je známý pod pojmem lineární trend, kde  $\beta_0$  představuje posun přímky po vertikální ose a  $\beta_1$  znázorňuje sklon přímky. Polynom druhého stupně (4.7) se označuje jako kvadratický trend. V grafu vizuálně symbolizuje parabolu. Polynom třetího stupně (4.8) se

nazývá kubickým trendem a v grafu znázorňuje kubickou parabolu. Rovnice (4.9) značí polynomický trend  $k$ -tého stupně. Obecně  $\beta_i$ ,  $i=1, 2, \dots, k$ , představují parametry funkce trendu a  $t^i$ ,  $i=1, 2, \dots, k$ , kde  $i$  znázorňuje exponent polynomu,  $k$  symbolizuje stupeň polynomu a  $t$  je časová složka.

Často se používá v časových řadách exponenciální a logaritmický trend. **Exponenciální trend** je vyjádřen (Cipra, 1986, s. 34)

$$T_t = \beta_0 \beta_1^t, \quad \beta_1 > 0. \quad (4.10)$$

Ve vzorci (4.10) představují  $\beta_0$  a  $\beta_1$  parametry, které jsou konstantní a jedinou proměnnou složkou je čas  $t$ . **Logaritmický trend** je popsán rovnicí (Fitzmaurice a kol., 2004, s. 14)

$$T_t = \beta_0 + \beta_1 \ln(t). \quad (4.11)$$

Ze vztahu (4.11) charakterizují  $\beta_0$  a  $\beta_1$  parametry funkce a proměnná  $t$  symbolizuje časovou složku.

Další trendové funkce představují tzv. S-křivky, které jsou specifické svým dvojitým esovitým prohnutím. Mezi ně se řadí například logistický a Gompertzův trend. Rozdíl mezi oběma trendy spočívá v tom, že **logistický trend** je symetrický a Gompertzův trend je asymetrický. Cipra (1986, s. 38) ve své knize uvádí předpis logistického trendu

$$T_t = \frac{\gamma}{1 - \alpha \beta^t}, \quad \beta > 0, \quad \gamma > 0. \quad (4.12)$$

Vzorec (4.12) obsahuje tři parametry  $\alpha$ ,  $\beta$  a  $\gamma$  a časovou proměnnou  $t$ . **Gompertzův trend** je dán rovnicí (Ord a Fildes, 2013, s. 296)

$$T_t = a e^{-b e^{-c t}}. \quad (4.13)$$

Funkce (4.13) zahrnují rovněž tři parametry  $a$ ,  $b$  a  $c$  a proměnnou  $t$ , která symbolizuje čas.

Výběr vhodné trendové funkce se nejčastěji určuje prostřednictvím vizuálního úsudku nebo vychází z vlastností časové řady. Správnému rozhodnutí napomáhají také statistické testy (t-testy, F-test), koeficient determinace a další. Problém ale spočívá v tom, že při identifikaci trendové funkce nelze nikdy zcela vyloučit subjektivní vlivy (Cipra, 1986, s. 29). Parametry trendové funkce se odhadnou prostřednictvím lineární nebo nelineární regrese.

## 4.2 Box-Jenkinsova metodologie

Box-Jenkinsova metodologie bude v této práci použita pro tvorbu dílčích prognóz úmrtnosti, plodnosti a migrace prostřednictvím Lee-Carterova modelu. Obecně zahrnuje mnoho modelových postupů, kterými lze analyzovat časové řady a vytvářet jejich predikce. Mezi nejznámější modelové postupy užívané k rozboru časových řad patří autoregresní proces (AR), proces klouzavých součtů (MA), ARMA smíšené procesy, ARIMA integrované procesy a další. Pojetí časové řady je na tomto místě ale trochu odlišné od předchozích přístupů. Box-Jenkinsova metoda vnímá časovou řadu jako soubor náhodných výkyvů, které mají stochastický charakter. Soustředí se na rozbor nesystematické části časové řady, která je charakterizována reziduální (náhodnou) složkou tvořenou vzájemně korelovanými náhodnými veličinami. Základem celého postupu je zkoumání závislostí prostřednictvím korelační analýzy. Dekompoziční metody se naopak

zaměřují na systematickou část časové řady, kterou tvoří trendová, sezónní a cyklická složka, kde jednotlivé složky jsou vzájemně nekorelované (Cipra, 1986, s. 19).

Jako každá metoda i Box-Jenkinsova metoda má své silné a slabé stránky. Mezi silné stránky nesporně patří, že se jedná o stochastické modely. Tím lze modelovat i takové časové řady, na které by jiné modely nestačily. Další výhodou je přesnost předpovědí. Vzhledem k porovnání středních čtvercových chyb dává Box-Jenkinsova metoda nejlepší výsledky v porovnání s běžnými metodami. Výhodou je rovněž odstranění subjektivních vlivů, kdy dochází k automatizaci celého procesu. Mezi slabé stránky patří velmi složitá až abstraktní interpretace výsledků. Je náročné vysvětlit běžnému uživateli, že lze vytvářet modely prostřednictvím náhodných šoků. Další slabou stránkou této metody je, že neumožňuje analyzovat krátké časové řady (Cipra, 1986, s. 100–101).

**Předpokladem pro užití Box-Jenkinsovy metody** je podle O'Donovana (1983, s. 32) minimální délka časové řady o 30 pozorování, ale shoduje se s Diekmannem a Mitterem (1984, s. 216), že pro statistickou stabilitu je vhodnější minimálně 50 pozorování. Před samotnou analýzou jednotlivých modelových postupů je nutné vymezit několik základních pojmů, mezi které patří náhodná veličina, stochastický proces, stacionarita, autokorelační funkce, parciální autokorelační funkce, proces bílého šumu a lineární proces. **Náhodná veličina**  $Y(\omega)$  přiřazuje každému elementárnímu jevu  $\omega$  konkrétní hodnotu  $y$  (Neubauer a kol., 2016, s. 89). Pro upřesnění tohoto pojmu lze uvést fiktivní příklad, kdy se sleduje počet sňatků v malé vesnici se třemi páry kohabituujících osob. Náhodná veličina  $Y$  označuje počet sňatků v této konkrétní vesnici. Elementární jevy  $\omega$  jsou vymezeny případy, které mohou nastat, tzn. žádný sňatek, jeden sňatek atd. K těmto jevům jsou přiřazeny konkrétní hodnoty 0, 1, 2, 3. Další příklady náhodných veličin uvádí Zvara a kol. (2013, s. 66–67). Problém ale nastává, pokud se počet těchto sňatků sleduje v čase  $t$ . Náhodná veličina počtu sňatků v této konkrétní vesnici v roce  $t$ , závisí na konkrétních hodnotách této veličiny v roce  $t - 1$ .

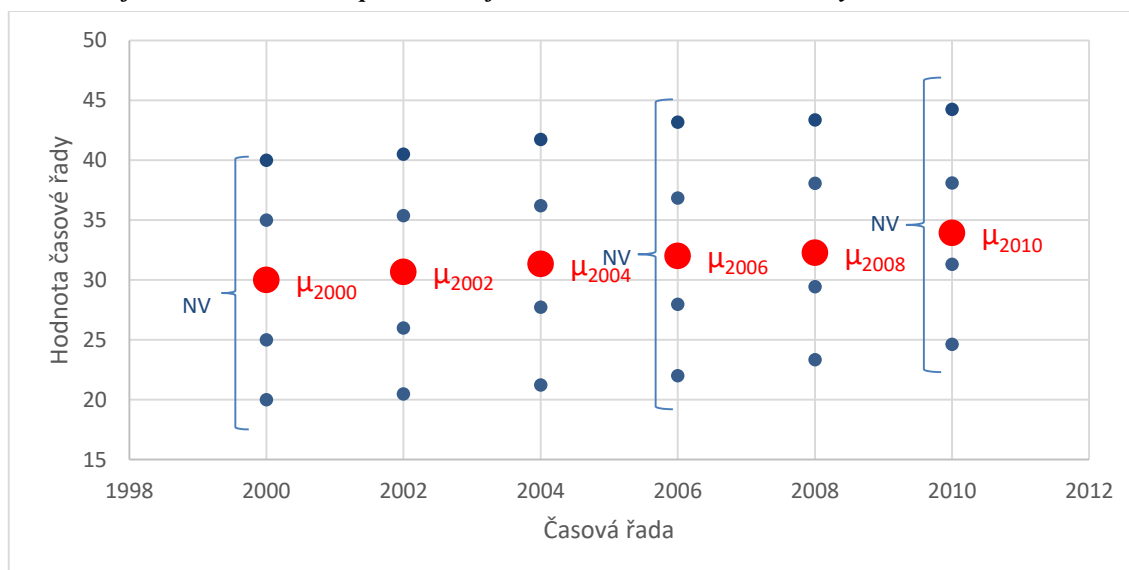
**Stochastický proces** je definovaný jako „(...) v čase uspořádaná řada náhodných veličin (...)“ (Arlt a Arltová, 2003, s. 30). Schematicky je znázorněn na obrázku 9. Stochastický proces lze popsat základními charakteristikami (Arlt a kol., 2002, s. 85; Cipra, 1986, s. 101)

$$\mu_t = E(y_t), \quad (4.14)$$

$$\sigma_t^2 = D(y_t) = E(y_t - \mu_t)^2, \quad (4.15)$$

$$\gamma(t, t - k) = C(y_t, y_{t-k}) = E(y_t - \mu_t)(y_{t-k} - \mu_{t-k}), \quad k \neq 0, \quad (4.16)$$

kde  $y_t$  představuje hodnotu časové řady v čase  $t$ ,  $y_{t-k}$  označuje hodnotu časové řady zpožděnou o  $k$  pozorování,  $E$  symbolizuje očekávanou střední hodnotu,  $D$  znázorňuje rozptyl,  $C$  kovarianci,  $\mu_t$  značí obecně střední hodnotu časové řady v čase  $t$ ,  $\sigma_t^2$  vyjadřuje obecně rozptyl v čase  $t$  a  $\gamma(t, t - k)$  ztělesňuje obecně autokovarianci mezi dvěma okamžiky v čase  $t$  a  $t - k$ . Na obrázku 9 je také zobrazena střední hodnota náhodné veličiny v čase  $t$ . Empiricky by ji bylo možné získat například jako průměr opakovaného měření teplot ve stejném čase. Problémem ale je, že k dispozici bývá často pouze jedno měření. Obecně je stochastický proces určen středními hodnotami v čase  $t$  (4.14), rozptyly v čase  $t$  (4.15) a kovariancemi mezi dvěma okamžiky v čase  $t$  a  $t - k$  (4.16).

**Obr. 9: Definice stochastického procesu na fiktivních hodnotách časové řady**

**Zdroj:** vlastní zpracování.

**Pozn.:** *NV* značí náhodnou veličinu a  $\mu$  představuje střední hodnotu.

Dalším důležitým pojmem je **stacionarita**. Jedná se o určité zjednodušení, které snižuje počet parametrů modelu. Rozlišuje se striktní stacionarita a její mírnější varianta, tzv. slabá stacionarita. Striktní stacionarita předpokládá, že „(...) pravděpodobnostní chování příslušného stochastického procesu je invariantní vůči posunům v čase (tj. pravděpodobnostní rozdělení náhodného vektoru  $(y_{t1}, \dots, y_{tk})$  je stejné jako rozdělení vektoru  $(y_{t1+h}, \dots, y_{tk+h})$  pro libovolné  $h$ ) (...)“ (Cipra, 1986, s. 101). Zjednodušeně lze napsat, že stochastický proces nepodléhá změnám v čase a chová se stále stejným způsobem (Arlt a Arltová, 2003, s. 31).

Jelikož se striktně stacionární procesy ve světě téměř nevyskytují, byla proto zavedena slabá stacionarita, která se omezuje pouze na střední hodnotu, rozptyl a kovarianci. Základní charakteristiky při uvažování slabé stacionarity lze popsat (Arlt a kol., 2002, s. 85)

$$\mu_t = \mu, \quad (4.17)$$

$$\sigma_t^2 = \sigma^2, \quad (4.18)$$

$$\gamma(t, t-k) = \gamma(t+k, t) = \gamma_k, \quad k \neq 0. \quad (4.19)$$

kde  $y_t$  vyjadřuje hodnotu časové řady v čase  $t$ ,  $y_{t-k}$  symbolizuje hodnotu časové řady zpožděnou o  $k$  pozorování,  $\mu_t$  je střední hodnota časové řady v čase  $t$ ,  $\sigma_t^2$  představuje rozptyl v čase  $t$  a  $\gamma(t, t-k)$  značí autokovarianci mezi dvěma okamžiky v čase  $t$  a  $t-k$ . Slabá stacionarita uvažuje neměnnou střední hodnotu  $\mu$  (4.17) nezávislou na čase  $t$ , neměnný rozptyl  $\sigma^2$  (4.18), který rovněž nezávisí na čase  $t$  a kovarianční funkce  $\gamma_k$  (4.19) je neměnná pro vzdálenost  $k$ , tzn. závislosti mezi náhodnými veličinami jsou shodné mezi po sobě následujícími roky ob jedno, ob dvě, ... (Arlt a Arltová, 2003, s. 31). **Box-Jenkinsova metoda umožňuje analyzovat pouze stacionární časové řady.** Pokud je časová řada nestacionární, lze ji velmi často na stacionární převést prostřednictvím transformací (Cipra, 1986, s. 101).

**Autokovarianční a autokorelační funkce** vyjadřují sílu lineární závislosti mezi hodnotou časové řady v čase  $t$  dále jen  $y_t$  a hodnotou časové řady zpožděnou o  $k$  označenou  $y_{t-k}$  (Arlt a kol.,

2002, s. 85). Rozdíl mezi autokovarianční a autokorelační funkcí spočívá v tom, že autokovarianční funkce nabývá bezrozměrných čísel, a aby mohla být lépe interpretována, je vhodné ji normovat na autokorelační funkci (Otyepka a kol., 2013, s. 12). **Parciální autokovarianční a autokorelační funkce** vycházejí z myšlenky, že síla lineární závislosti mezi  $y_t$  a  $y_{t-k}$  může být způsobena vztahem s dalšími veličinami  $y_{t-1}, \dots, y_{t-k+1}$ . Parciální autokorelační funkce „(...) podává informaci o korelaci veličin  $y_t$  a  $y_{t-k}$  očištěnou o vliv veličin ležících mezi nimi (...)“ (Arlt a kol., 2002, s. 85). Graficky lze autokorelační funkci a parciální autokovarianční funkci vyjádřit korelogramem, který se používá pro zhodnocení reziduí, pokud jsou procesem bílého šumu, a také se uplatňuje při identifikaci volby procesu Box-Jenkinsovy metodologie (Arltová a Arlt, 2009, s. 490).

Doposud byly uvažovány základní charakteristiky v teoretické rovině. Za předpokladu stacionarity souboru lze vytvořit **odhad základních charakteristik** (Arlt a kol., 2002, s. 86; Cipra, 1986, s. 101)

$$\bar{y} = \sum_{t=1}^n \frac{y_t}{n}, \quad (4.20)$$

$$s^2 = c_0 = \frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n}, \quad (4.21)$$

$$c_k = \sum_{t=1}^{n-k} \frac{(y_t - \bar{y})(y_{t-k} - \bar{y})}{n}, \quad k = 0, 1, \dots, n-1, \quad (4.22)$$

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, n-1, \quad (4.23)$$

kde  $y_t$  značí hodnotu časové řady v čase  $t$ ,  $y_{t-k}$  představuje hodnotu časové řady zpožděnou o  $k$  pozorování,  $\bar{y}$  symbolizuje výběrový průměr hodnot časové řady,  $n$  je počet pozorování,  $c_0$  a  $s^2$  vyjadřují výběrový rozptyl hodnot časové řady,  $c_k$  znázorňuje výběrovou autokovarianční funkci pro vzdálenost  $k$ ,  $r_k$  představuje výběrovou autokorelační funkci pro vzdálenost  $k$ . Podle vzorce (4.20) je odhadem střední hodnoty  $\mu$  výběrový průměr, odhadem rozptylu  $\sigma$  je výběrový rozptyl (4.21), kde výběrový rozptyl je speciální případ výběrové autokovariance (4.22) při  $k=0$ . Odhadem autokovarianční funkce  $\gamma$  je výběrová autokovarianční funkce (4.22) a odhadem autokorelační funkce  $\rho$  je výběrová autokorelační funkce (4.23), která se vypočítá jako podíl výběrové autokovarianční funkce a výběrového rozptylu.

Dalším pojmem je **proces bílého šumu**, který představuje nesystematickou složku časové řady a je základním prvkem všech lineárních procesů (Arlt a Arltová, 2003, s. 34). Proces bílého šumu byl popsán v podkapitole 4.1. Posledním pojmem je **lineární proces**. Tento proces je klíčový především proto, že z něj vycházejí všechny procesy Box-Jenkinsonovy metodologie. Praktické užití ale nemá (Cipra, 1986, s. 108). Na lineárním procesu se v této práci představí jiný způsob zápisu vzorců prostřednictvím operátoru zpětného zpoždění, který bude dále využit u složitějších modelů ARMA a ARIMA. Dále se popíše podmínka existence lineárního procesu a bude definována podmínka invertibility. Pro snadnější orientaci ve výpočtech je v tabulce 4 uveden seznam použitých symbolů. Nebude-li uvedeno jinak, bude toto značení použito pro popis lineárního procesu a dále pro podkapitoly 4.2.1 a 4.2.2.



Tab. 4: Seznam symbolů použitých u Box-Jenkinsovy metody

Symbol	Popis
$t$	čas
$y_t$	hodnota časové řady v čase $t$
$y_{t-j}$	hodnota časové řady, která je o $j$ kroků zpožděná, tzn., pokud $j = 1$ je časová řada zpožděna o jeden krok v čase $t - 1$
$\varepsilon_t$	hodnota procesu bílého šumu
$\varepsilon_{t-j}$	hodnota procesu bílého šumu, která je o $j$ kroků zpožděná, tzn., pokud $j = 1$ je hodnota procesu bílého šumu, která je zpožděna o jeden krok v čase $t - 1$
$B$	operátor zpětného zpoždění
$B^j$	operátor zpětného zpoždění řádu $j$
$\psi_j, \pi_j$	parametr lineárního procesu
$\varphi_j$	parametr autoregresního procesu
$\theta_j$	parametr procesu klouzavých součtů
$\delta$	parametr smíšeného procesu ARMA
$\psi(B), \pi(B)$	operátor lineárního procesu
$\varphi(B)$	autoregresní operátor
$\theta(B)$	operátor klouzavých součtů
$\Delta$	diferenční operátor
$d$	řád difference
$\varphi(B)\Delta^d$	zobecněný autoregresní operátor

**Zdroj:** Cipra (1986, s. 107–121), Cipra (1986, s. 141–143), vlastní zpracování.

Lineární proces je takový proces, který lze popsat rovnicí (Arltová a Arlt, 2009, s. 37)

$$y_t = \psi_0 + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots, \quad (4.24)$$

nebo za určitých podmínek rovnicí (Cipra, 1986, s. 107, 121)

$$y_t = \pi_0 + \pi_1 y_{t-1} + \pi_2 y_{t-2} + \dots + \varepsilon_t. \quad (4.25)$$

Jedná se o nekonečnou řadu zpožděných procesů bílého šumu (4.24) nebo za určitých okolností ekvivalentě o nekonečnou řadu zpožděných hodnot časové řady (4.25). Za podmínky nulové střední hodnoty časové řady  $y_t$  se první člen  $\psi_0$  a  $\pi_0$  z rovnice (4.24) a (4.25) vynechává. Matematicky je celý postup zpracován v příloze 2. Obě rovnice lze poté vyjádřit prostřednictvím operátoru zpětného zpoždění  $B$  (Cipra, 1986, s. 107)

$$y_t = \psi(B) \varepsilon_t, \quad (4.26)$$

$$\pi(B) y_t = \varepsilon_t, \quad (4.27)$$

kde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = 1 + \sum_{j=1}^{\infty} \psi_j B^j, \quad (4.28)$$

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = 1 - \sum_{j=1}^{\infty} \pi_j B^j. \quad (4.29)$$

Celé odvození zápisu prostřednictvím operátoru zpětného zpoždění je matematicky zpracováno v příloze 3. Jedná se pouze o jinou formu zápisu lineárního procesu, která se často užívá při složitějších záznamech Box-Jenkinsových procesů. Obě rovnice (4.26) a (4.27) tedy rovněž

vyjadřují lineární proces, kde operátor lineárního procesu  $\psi(B)$  je vyjádřen formulací (4.28) a operátor lineárního procesu  $\pi(B)$  je charakterizován vztahem (4.29).

Existence lineárního procesu a vztah mezi jeho operátory  $\psi(B)$  a  $\pi(B)$  je určen (Cipra, 1986, s. 107–108)

$$\psi(B) \text{ konverguje pro } |B| \leq 1, \quad (4.30)$$

$$\pi(B) \text{ konverguje pro } |B| \leq 1, \quad (4.31)$$

$$\pi(B) = \psi^{-1}(B). \quad (4.32)$$

**Podmínka pro existenci lineárního procesu** je definována formulací (4.30), která zabezpečuje zároveň stacionaritu lineárního procesu (4.24). Při platnosti **postačující podmínky pro invertibilitu** (4.31) lze vyjádřit lineární proces (4.24) také prostřednictvím rovnice (4.25), poté se tento proces nazývá invertibilní (Cipra, 1986, s. 108). Protože v reálných situacích nejsou k dispozici zpožděné hodnoty procesu bílého šumu, ale pouze zpožděné hodnoty časové řady, je nutné je odhadnout. Právě k tomuto účelu byl vytvořen vztah (4.32), kdy za podmínky invertibility se zpožděné hodnoty procesu bílého šumu odhadnou ze zpožděných hodnot časové řady (Martinková, 2012, s. 18).

#### 4.2.1. Vybrané stacionární procesy

Doposud byl popsán pouze obecný lineární proces, který představuje nekonečnou řadu zpožděných hodnot časové řady nebo zpožděných procesů bílého šumu. V praxi se ale používají pouze jeho obměny s konečným počtem zpoždění (Cipra, 1986, s. 108). Tato podkapitola se zabývá pouze těmi procesy, které jsou stacionární. Mezi ně se řadí autoregresní proces, proces klouzavých součtů a smíšené procesy ARMA.

Prvním z nich je **autoregresní proces (AR)**. Tento proces je charakteristický tím, že časová řada závisí na svých minulých pozorování. Autoregresní proces prvního řádu označovaný AR(1) lze zapsat (Cipra, 1986, s. 115)

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varepsilon_t. \quad (4.33)$$

Základní myšlenkou vzorce (4.33) je zahrnutí jednoho minulého pozorování do časové řady. Autoregresní proces druhého řádu AR(2) je určen (Cipra, 1986, s. 118)

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varepsilon_t. \quad (4.34)$$

Poslední hodnota časové řady je podle rovnice (4.24) vysvětlena dvěma jejími zpožděními. Obecně lze autoregresní proces řádu  $p$  pod označením AR(p) popsat (Ben-Daya a kol., 2009, s. 171; Prášková, 2016, s. 63)

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t, \quad (4.35)$$

nebo prostřednictvím parametru zpětného zpoždění za předpokladu nulové střední hodnoty časové řady lze ekvivalentně psát (Box a Pierce, 1970, s. 1512)

$$\varphi(B)y_t = \varepsilon_t. \quad (4.36)$$

Význam rovnice (4.35) a formulace (4.36) spočívá v zahrnutí konečného počtu zpožděných hodnot časové řady do autoregresního procesu.

**Proces klouzavých součtů (MA)** se vyznačuje tím, že se zahrnují do časové řady zpožděné náhodné šoky. Proces klouzavých součtů prvního řádu pod označením MA(1) se zapíše (Durbin, 1960, s. 239)

$$y_t = \varepsilon_t + \theta_0 + \theta_1 \varepsilon_{t-1}. \quad (4.37)$$

Principem vzorce (4.37) je zakomponování jedné zpožděné hodnoty náhodného šoku do procesu klouzavých součtů prvního řádu. Proces klouzavých součtů druhého řádu MA(2) se vypočítá (Cipra, 1986, s. 111)

$$y_t = \varepsilon_t + \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}. \quad (4.38)$$

Podstatou rovnice (4.38) je zařazení dvou zpožděných náhodných šoků do časové řady. Proces klouzavých součtů řádu  $q$  označený MA( $q$ ) lze obecně napsat (Durbin, 1960, s. 239; Ben-Daya a kol., 2009, s. 171)

$$y_t = \varepsilon_t + \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (4.39)$$

nebo prostřednictvím parametru zpětného zpoždění se za předpokladu nulové střední hodnoty časové řady vyjádří (Cipra, 1986, s. 109)

$$y_t = \theta(B) \varepsilon_t. \quad (4.40)$$

Proces klouzavých součtů (4.39) a (4.40) udává, kolik zpožděných náhodných šoků se zařadí do rovnice.

**Smíšený proces ARMA** se skládá z procesů AR a MA, tzn., časová řada závisí na svých předchozích pozorování a zároveň na zpožděných náhodných šocích. Smíšený proces ARMA(1,1) se vyjádří (Cipra, 1986, s. 119)

$$y_t = \delta + \varphi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}. \quad (4.41)$$

Základní myšlenkou rovnice (4.41) je zařazení jedné zpožděné hodnoty časové řady a jedné zpožděné hodnoty náhodného šoku do smíšeného procesu ARMA. Obecně se smíšený proces ARMA( $p, q$ ) zapíše (Durbin, 1960, s. 239; Prášková, 2016, s. 71)

$$y_t = \delta + \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (4.42)$$

nebo prostřednictvím parametru zpětného zpoždění za předpokladu nulové střední hodnoty časové řady lze napsat (Cipra, 1986, s. 119)

$$\varphi(B) y_t = \theta(B) \varepsilon_t. \quad (4.43)$$

Smíšený proces ARMA řádu  $p$  a řádu  $q$  (4.42) nebo (4.43) udává, kolik zpožděných hodnot časové řady a kolik zpožděných náhodných šoků se zařadí do rovnice.

#### 4.2.2. Nestacionární proces ARIMA

Příkladem nestacionárního procesu je ARIMA model, který se skládá z autoregresního procesu AR, procesu klouzavých průměrů MA a difference I. Za předpokladu nulové střední hodnoty časové řady se obecně model ARIMA( $p, d, q$ ) zaznamená (Box a Pierce, 1970, s. 1509)

$$\varphi(B) \Delta^d y_t = \theta(B) \varepsilon_t, \quad (4.44)$$

kde (Box a Pierce, 1970, s. 1509)

$$\Delta = 1 - B. \quad (4.45)$$

Smíšený integrovaný proces ARIMA řádu  $p$ ,  $d$  a  $q$  (4.44) představuje, kolik zpožděných hodnot časové řady, kolik zpožděných náhodných šoků a kolik řádů diferencí (4.45) se zařadí do modelu.

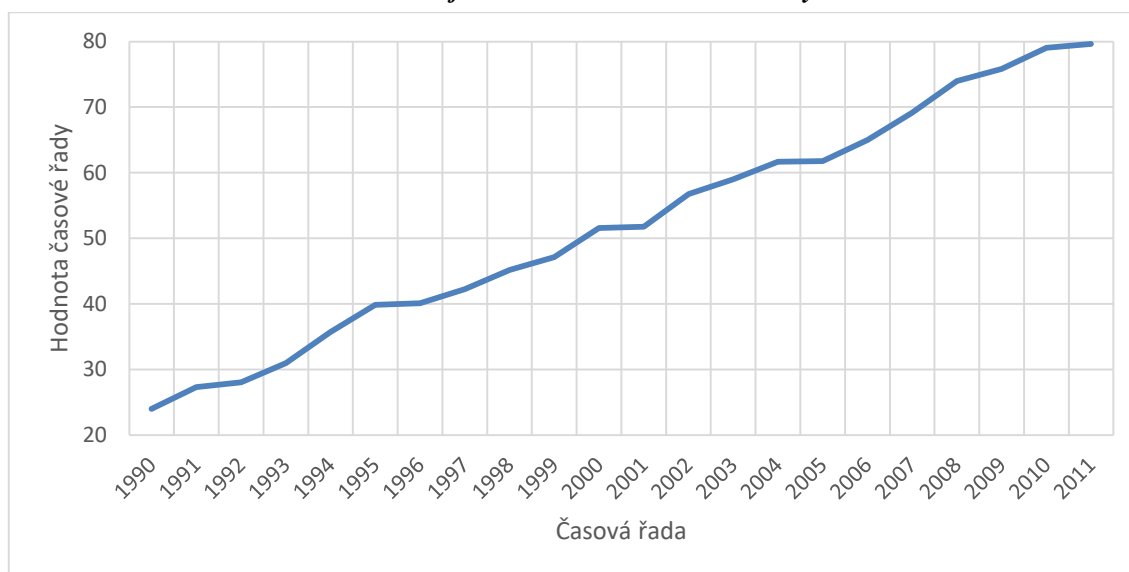
Diferencování se používá v případě, že časová řada není stacionární a obsahuje například trend, který byl vysvětlen v podkapitole 4.1. Prostřednictvím diferencování lze na stacionární časovou řadu převést (Cipra, 1986, s. 142)

$$(1-B)y_t = y_t - y_{t-1}, \quad (4.46)$$

$$(1-B)^2 y_t = (1-2B+B^2)y_t = y_t - 2y_{t-1} + y_{t-2}. \quad (4.47)$$

Pro ilustraci je uvedena první diference (4.46), prostřednictvím které je možné odstranit lineární trend (4.6) časové řady a druhá diference (4.47), na základě které se odstraní kvadratický trend (4.7) časové řady (Arlt a kol., 2002, s. 26).

**Obr. 10: Simulace lineárního trendu na fiktivních hodnotách časové řady**



**Zdroj:** vlastní zpracování.

Na obrázku 10 je simulována časová řada, která vykazuje přibližně lineární trend. Při analýze časové řady prostřednictvím trendové funkce by se nejprve odhadly parametry lineárního trendu a odečtením skutečných hodnot od hodnot trendu by se poté odstranil lineární trend. Při použití Box-Jenkinsovy metodologie je postup trochu odlišný a k odstranění trendu se využívá diferencování. Pro odstranění lineárního trendu se používá prvních diferencí. Na obrázku 11 je znázorněna modrou barvou časová řada po odstranění lineárního trendu prostřednictvím prvních diferencí, která kolísá kolem nenulové hodnoty. Vzhledem k tomu, že není splněn předpoklad nulové střední hodnoty časové řady, je nutné diferencovanou časovou řadu centrovat, tzn. od hodnot diferencované časové řady se odečte jejich průměr (Cipra, 1986, s. 121). Výsledek takto upravené časové řady je zobrazen na obrázku 11 červenou barvou. Takto zpracovanou časovou řadu tvoří náhodné šoky, které jsou dále modelovány prostřednictvím procesů AR, MA nebo jejich kombinací ARMA. Obecně lze výpočet smíšeného procesu ARIMA bez předpokladu nulové střední hodnoty vyjádřit rovnicí (Hyndman a Athanasopoulos, 2017)

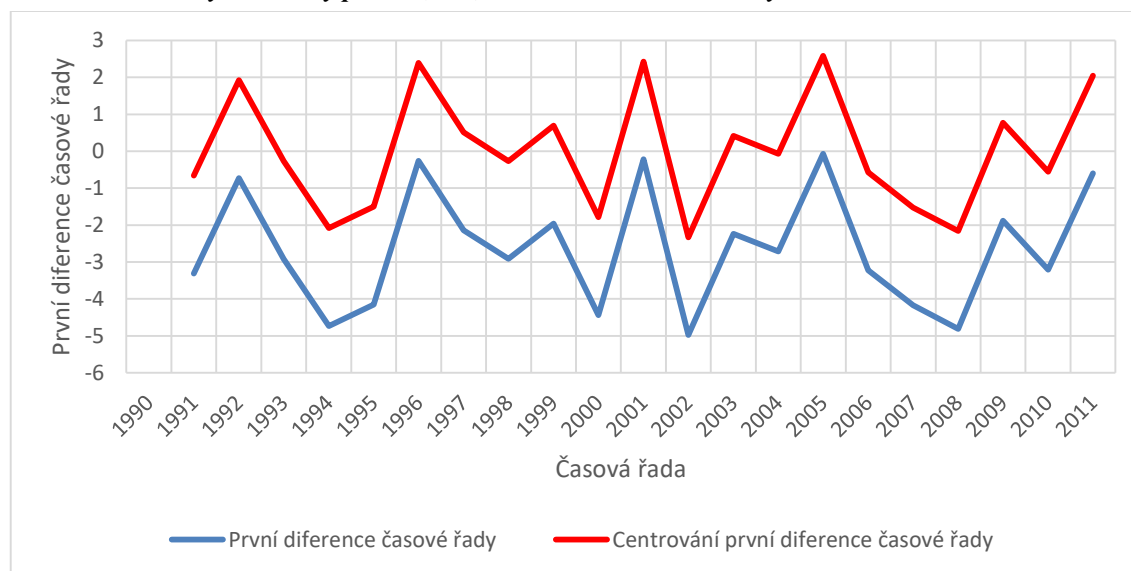
$$\varphi(B)\Delta^d y_t = c + \theta(B)\varepsilon_t, \quad (4.48)$$

eventuelně (Hyndman a Athanasopoulos, 2017)

$$\varphi(B)\Delta^d \left(y_t - \frac{\mu t^d}{d!}\right) = \theta(B)\varepsilon_t, \quad (4.49)$$

kde  $c$  představuje konstantu,  $\mu$  značí střední hodnotu  $\Delta^d y_t$ ,  $t$  symbolizuje čas a  $d!$  je faktoriál řádu difference. Cílem rovnic (4.48) a (4.49) je stacionarizovat původní časovou řadu a odhadnout její budoucí vývoj. Hyndman a Athanasopoulos (2017) nazývají střední hodnotu  $\mu$  driftem, pokud je faktoriál řádu difference jedna.

**Obr. 11:** Odstranění trendu prostřednictvím první difference a následné centrování nově vzniklých hodnot časové řady na nulový průměr, tzv., stacionarizace časové řady



**Zdroj:** vlastní zpracování.

## Kapitola 5

### Východiska pro zpracování regionální prognózy obyvatel Jihočeského kraje

Tato kapitola pojednává o sběru datových zdrojů a o problémech jejich použití pro prognózu obyvatel. Základní otázkou je vymezení území, pro které bude prognóza zpracována. K tomuto účelu bylo zvoleno území Jihočeského kraje s cílem vytvoření regionální prognózy. Při studiu jakéhokoli územního celku je nutné sledovat jeho územní změny v čase. Tomuto tématu se bude detailně věnovat podkapitola legislativních změn ve vymezení Jihočeského kraje. Dále je nutné zabývat se otázkou sběru a kvality dat. Datové zdroje pro území Jihočeského kraje poskytuje Český statistický úřad (ČSÚ, 2017a). Cílem této práce je v závislosti na typu dat navrhnout metodické zpracování populační prognózy Jihočeského kraje. V této souvislosti se nabízí dvojitý přístup ke sběru dat. První z nich má za cíl simulovat špatně dosažitelné, méně kvalitní a relativně za krátký horizont dostupné údaje. Pro další analýzu bude tato řada nazvána jako **krátká výchozí časová řada**. Druhý přístup má za cíl simulovat velmi kvalitní a za poměrně dlouhé časové období dostupná data. Pro zjednodušení se tato řada bude dále nazývat **dlouhá výchozí časová řada**. Odlišný způsob sběru dat k oběma přístupům bude dále popsán v podkapitole datové zdroje.

#### 5.1 Legislativní změny ve vymezení území Jihočeského kraje

Zákon o územním členění státu z roku 1960 vytvořil na území České republiky sedm krajů a na území Slovenska kraje tři. Samostatnou územní jednotku tvořilo podle tohoto zákona hlavní město Praha. Legislativně bylo definováno, že kraje se formují z okresů a ty jsou tvořeny obcemi. U hlavního města Prahy bylo toto členění trochu odlišné. Skládalo se, na rozdíl od klasifikace krajů, z deseti obvodů. Legislativa vymezila i území Jihočeského kraje, které se skládalo z okresů České Budějovice, Český Krumlov, Jindřichův Hradec, Pelhřimov, Písek, Prachatice, Strakonice a Tábor. Zákon o územním členění státu byl několikrát novelizován, ale doposud nebyl zrušen (Československo, 1960, s. 98).

Administrativní vymezení obcí a krajů je zakotveno v Ústavě České republiky, kde obce jsou základními územními samosprávnými celky a kraje jsou vyššími územními samosprávnými celky (Česko, 1992, s. 15). V roce 1997 byl přijat klíčový ústavní zákon o vytvoření vyšších územních samosprávných celků, který zřídil 13 krajů a hlavní město Prahu, tedy 14 vyšších územních

samosprávných celků. Účinnost tohoto zákona nastala 1. 1. 2000. Budějovický kraj se od tohoto roku skládá z okresů České Budějovice, Český Krumlov, Jindřichův Hradec, Písek, Prachatice, Strakonice a Tábor (Česko, 1997, s. 7018; Kadeřábková a Peková, 2012, s. 27).

V roce 2001 nabývá účinnosti nový ústavní zákon č. 176/2001 Sb., kterým se mění ústavní zákon č. 347/1997 Sb. a dochází k přejmenování 4 krajů. Jihlavský kraj se od tohoto data označil jako Vysočina, Brněnský kraj byl pojmenován Jihomoravským krajem, Ostravský kraj byl nazván Moravskoslezským krajem a Budějovický kraj byl přejmenován na Jihočeský kraj (Česko, 2001, s. 3674). V roce 2011 dochází k další změně ústavního zákona č. 347/1997. Tato novela ale nemá vliv na vymezení Jihočeského kraje. Upravuje pouze název Vysočina a nahrazuje ho pojmem Kraj Vysočina (Česko, 2011, s. 1314).

## 5.2 Datové zdroje využité v analytické části práce a jejich přímá souvislost s prahem a horizontem prognózy

Datové zdroje patří k limitujícím faktorům každé analýzy. Přesto i s malým množstvím výchozích dat lze vytvářet poměrně kvalitní analýzy. Statisticky je ale nutné počítat s větší chybou. Přesnější analýzy lze zpracovávat s větším množstvím dat. Délka výchozí časové řady má pro analýzu značný význam. Cílem této práce je ilustrovat tuto skutečnost na datech. Z tohoto důvodu probíhal sběr dat ve dvou etapách, zvlášť pro krátkou výchozí časovou řadu a zvlášť pro dlouhou výchozí časovou řadu.

**Krátká výchozí časová řada** reprezentuje údaje za Jihočeský kraj, které jsou dostupné ze stránek Českého statistického úřadu. V rámci úmrtnosti jsou k dispozici dvouleté úmrtnostní tabulky podle jednotek věku od roku 2002–2003 do roku 2014–2015 za ženy (ČSÚ, 2016b; ČSÚ, 2015b; ČSÚ, 2014b; ČSÚ, 2013b; ČSÚ, 2012b; ČSÚ, 2011b; ČSÚ, 2010b; ČSÚ, 2009b; ČSÚ, 2008b; ČSÚ, 2007b; ČSÚ, 2006b; ČSÚ, 2005b; ČSÚ, 2004b) a muže (ČSÚ, 2016a; ČSÚ, 2015a; ČSÚ, 2014a; ČSÚ, 2013a; ČSÚ, 2012a; ČSÚ, 2011a; ČSÚ, 2010a; ČSÚ, 2009a; ČSÚ, 2008a; ČSÚ, 2007a; ČSÚ, 2006a; ČSÚ, 2005a; ČSÚ, 2004a). Míry plodnosti podle jednotek věku je možné dohledat již od roku 1986, kdy už by se jednalo o dlouhou časovou řadu. Problém ale nastává v legislativě, kdy 1. 1. 2000 nabývá účinnosti ústavní zákon o vytvoření vyšších územních samosprávných celků. Tento zákon mění území Budějovického kraje o okres Pelhřimov, který se stává součástí kraje Vysočina (Česko, 1997, s. 7018). Pro populační prognózu musí být území Jihočeského kraje stejně územně vymezené. Z tohoto důvodu lze použít časovou řadu měr plodnosti podle věku z Českého statistického úřadu pouze od roku 2000 (ČSÚ, 2016c; ČSÚ, 2010c; ČSÚ, 2008c). Počáteční stavy žen (ČSÚ, 2016e; ČSÚ, 2015d; ČSÚ, 2014d; ČSÚ, 2013d; ČSÚ, 2012d; ČSÚ, 2011d; ČSÚ, 2010e; ČSÚ, 2009d; ČSÚ, 2008d; ČSÚ, 2007c; ČSÚ, 2006c; ČSÚ, 2005c; ČSÚ, 2004c) a mužů (ČSÚ, 2016d; ČSÚ, 2015c; ČSÚ, 2014c; ČSÚ, 2013c; ČSÚ, 2012c; ČSÚ, 2011c; ČSÚ, 2010d; ČSÚ, 2009c; ČSÚ, 2008d; ČSÚ, 2007c; ČSÚ, 2006c; ČSÚ, 2005c; ČSÚ, 2004c) podle jednotek věku jsou k dispozici ze stránek ČSÚ od 1. 1. 2003 pro Jihočeský kraj. Data za migraci podle jednotek věku nejsou ze stránek ČSÚ k dispozici, proto je bude nutné dopočítat. Tento problém bude popsán v podkapitole 6.4.

Obecně platí, že délka prognózy od prahu do jeho horizontu by měla vycházet z délky výchozí časové řady. Na základě této myšlenky je délka prognózy z krátké výchozí časové řady stanovena na desetileté období, tzn. jedná se o krátkodobou prognózu. **Prahem** této prognózy je 1. 1. 2015 a jejím **horizontem** je 1. 1. 2025.

**Za dlouhou výchozí časovou řadu** jsou považována data poskytnutá Českým statistickým úřadem pro účel této diplomové práce. K dispozici jsou živě narození a zemřelí podle jednotek věku a pohlaví od roku 1986 do roku 2015. Dále jsou užity koncové stavy obyvatel podle jednotek věku od roku 1985 do roku 2015. Rovněž jsou k dispozici data za vystěhovalé a přistěhovalé muže a ženy podle jednotek věku od roku 1993 do roku 2015. Na základě této výchozí časové řady byla stanovena délka prognózy na dvacetileté období, tzn. jedná se o střednědobou prognózu. **Prahem** této prognózy je 31. 12. 2015 a jejím **horizontem** je 31. 12. 2035.



## Kapitola 6

### Prognóza vývoje počtu obyvatel Jihočeského kraje prostřednictvím analyzování vývoje krátké výchozí časové řady

Tato kapitola se zaměřuje na problematiku krátkých časových řad. Budou zde aplikovány vybrané demografické modely a statistické postupy v analýze úmrtnosti, plodnosti a migrace na krátkou výchozí časovou řadu s cílem vytvoření co nejadekvátnější populační prognózy Jihočeského kraje do počátku roku 2015. Pro dílčí úmrtnosti byl zvolen Heligman-Pollard osmiparametrový model z důvodu, protože modeluje pravděpodobnosti úmrtí ve všech věcích. Při analýze krátké časové řady u plodnosti byla vybrána Beta funkce, která modeluje míry plodnosti rovněž v každém věku. Výhodou obou modelů je také to, že časovou řadu vyrovnají. Pro dílčí prognózu migrace byla zvolena prognóza prostřednictvím dolního kvartilu, mediánu a horního kvartilu migračního salda podle jednotek věku v čase. Celá problematika včetně nepřímé metody výpočtu migračního salda bude popsána v podkapitole 6.3.

Ke zpracování populační prognózy byly zvoleny programy Microsoft Excel, MORTPAK (United Nations, 2013b), SAS (SAS Institute Inc, 2017) a program R (R Core Team, 2016). Microsoft Excel 2016 se stal široce využívaným tabulkovým procesorem. MORTPAK je volně dostupný program ke stažení, užívaný převážně pro demografické modelování v rozvojových zemích (United Nations, 2013a, s. iii). Program SAS je široce používaný program, který se používá jak při demografických, tak i ve statistických výpočtech a rovněž podporuje databázové zpracování. Program R je open source program,<sup>4</sup> který se užívá převážně pro statistiku. Tento program je také znám díky balíčkům (angl. packages). Jedná se o balíčky, které jsou vypracované dobrovolníky případně lidmi z týmu kolem tohoto programu. Ty se poté doinstalovávají do programu R a usnadňují tak práci s psaním kódu.

---

<sup>4</sup> **Open source** lze přeložit jako otevřený zdroj, kdy každý může číst a upravovat zdrojový kód, ale stejně tak je povinen uveřejnit pod touto licencí svůj program (Roubal, 2010, s. 79).

## 6.1 Dílčí prognóza úmrtnosti s využitím modelu Heligmana-Pollarda

Pro vytvoření dílčí prognózy úmrtnosti z krátké výchozí časové řady byl zvolen osmiparametrový model Heligmana-Pollarda. Výběr modelu je zapříčiněn jeho nespornou výhodou, že modeluje kvocienty úmrtnosti v každém věku. Jak již bylo řečeno úmrtnostní tabulky za muže a za ženy jsou k dispozici z Českého statistického úřadu vždy za dvouletý interval od roku 2002–2003 do roku 2014–2015. Dvouleté tabulky úmrtnosti se počítají z důvodu zpřesnění dat, jelikož za malé územní celky je méně úmrtí a kvocienty úmrtnosti podle věku by byly značně rozkolísané. Kvůli zachování konzistentnosti dat bude i budoucí odhad těchto kvocientů vždy za dvouleté období.

K odhadu osmi parametrů Heligmana-Pollarda se využije programu MORTPAK (United Nations, 2013b). Podrobný postup je popsán v manuálu United Nations (2013a, s. 67–68). Výhoda tohoto programu spočívá v rychlém výpočtu a snadné ovladatelnosti programu. Naopak nevýhodou je odhad z  $n$ -letých kvocientů úmrtnosti. Parametry jsou kvůli tomu zatíženy chybou. Pro zmírnění této chyby se výpočet zkrácených úmrtnostních tabulek provádí z úplných úmrtnostních tabulek zvlášť pro muže a pro ženy. Nejprve se z úplné úmrtnostní tabulky vyberou tabulkové počty dožívajících (značené  $l_x$ ) v přesném věku  $0$ ,  $1$  a dále v přesném věku  $5$ ,  $10$ , ...,  $100$ , vždy po pěti letech. Dále se postupuje podle klasického výpočtu zkrácených úmrtnostních tabulek, jak uvádí Pavlík a kol. (1986, s. 185)

$${}_nd_x = l_x - l_{x+n}, \quad x = 0 \wedge n = 1, \quad x = 1 \wedge n = 4, \quad x = 5, \dots, 100 + \wedge n = 5, \quad (6.1)$$

$${}_nq_x = \frac{{}_nd_x}{l_x}, \quad x = 0 \wedge n = 1, \quad x = 1 \wedge n = 4, \quad x = 5, \dots, 100 + \wedge n = 5, \quad (6.2)$$

kde  $n$  je šířka intervalu a  $x$  je věk. Podle vzorce (6.1) se vypočítají tabulkové počty zemřelých v dokončeném věku  $x$  až  $x + n$  označené  ${}_nd_x$  a po dosazení do vzorce (6.2) se získají kvocienty úmrtnosti mezi přesnými věky  $x$  a  $x + n$  symbolizované  ${}_nq_x$  (Pavlík a kol., 1986, s. 726).

Jednotlivé parametry modelu Heligmana-Pollarda lze prognózovat do budoucna. Problémem ale je, že mezi těmito parametry mohou být vzájemné vztahy a z tohoto důvodu nelze modelovat každý parametr zvlášť. Řešení nabízí například Keilman a kol. (2001, s. 53), který modeluje budoucí prognózu parametrů Heligmana-Pollarda prostřednictvím vícerozměrného modelu MARIMA (multivariate ARIMA). Tento postup se ukázal pro tuto práci nevhodný ze dvou důvodů. Prvním důvodem jsou ne zcela kvalitní odhady parametrů Heligmana-Pollarda způsobené agregací dat na pětileté věkové intervaly. Druhým důvodem je velmi krátká výchozí časová řada, kdy se promítají do prognózy prostřednictvím zpoždění špatné odhady parametrů, které způsobují výrazné výkyvy.

Pro identifikaci vazeb mezi parametry Heligmana-Pollarda byla v této práci zvolena faktorová analýza, která se tradičně aplikuje jak v českém, tak i v zahraničním prostředí. Příkladem jsou studie Hulíka a kol. (2011, s. 13–34) nebo Bircha a kol. (2001, s. 201–210). Z hlediska matematického postupu popisuje faktorovou analýzu Matyašovská (2014, s. 30–36). V této práci je zpracována v programu SAS (SAS Institute Inc, 2017). Po volbě vhodného počtu faktorů na základě vlastních čísel se vypočítají faktorová skóre, která se extrapolují do budoucna. Pro tuto extrapolaci jsou zvoleny trendové funkce, které jsou vhodné i pro kratší časové řady.

Následně se vyčíslí vztah mezi faktorovými skóry a původními hodnotami parametrů Heligmana-Pollarda, které je vhodné upravit na z-skóry tzn. odečíst průměr a vydělit směrodatnou odchylkou (Zvára a kol., 2013, s. 22). Úprava na z-skóry se provádí za účelem korekce na stejné měřítko, kdy různé měřítko v regresní analýze může způsobit výrazné vychýlení dané regresní křivky. Samotný vztah se poté vyčíslí jednoduchou lineární regresí, kterou popisují například Montgomery a kol. (2012, s. 12), kdy za  $x$  se dosadí faktorová skóre daného faktoru a za  $y$  příslušné parametry Heligmana-Pollarda upravenými na z-skóry. V tomto případě vychází nulový průsečík přímky s osou  $y$  způsobený úpravou na z-skóry a sklon přímky je totožný s rotovaným řešením faktorové analýzy. Do tohoto lineárního vztahu se dosadí faktorová skóre odhadnutá prostřednictvím trendové funkce a výsledkem je očekávaný vývoj parametrů Heligmana-Pollarda, které jsou stále ještě upravené na z-skóry. Z tohoto důvodu je nutné vynásobit tyto parametry příslušnou směrodatnou odchylkou a přičíst průměr. Takto zpracované parametry se dosadí do původní rovnice Heligmana-Pollarda a vypočítají se kvocienty úmrtnosti.

Vzhledem k velkému množství vztahů mezi parametry Heligmana-Pollarda byl navržen specifický přístup k řešení intervalů spolehlivosti a to prostřednictvím simulace. V principu se hodnoty faktorového skóre nemusí pohybovat přímo na hranici zvoleného trendu, ale měly by být vzdáleny co nejblíže k tomuto trendu. Toto průměrné kolísání kolem trendu se měří variabilitou. Vzhledem ke skutečnosti, že téměř ve všech případech bude použit jako nejvhodnější logaritmický trend, lze pro předpověď variability použít metodologii jednoduché lineární regrese popsanou například v knize autorů Yana a Xiaogang Sua (2009, s. 20). Samotná simulace poté spočívá v generování náhodných čísel z normálního rozdělení se střední hodnotou popsanou trendem daného parametru Heligmana-Pollarda a směrodatnou odchylkou, která se vypočítá jako odmocnina z předpovědi variability.

Pro simulaci bylo zvoleno 10 000 opakování zvláště pro každý rok. Následně jsou prostřednictvím simulovaných parametrů Heligmana-Pollarda vypočítány kvocienty úmrtnosti. Z nově vzniklých kvocientů úmrtnosti je vypočítán pro každý rok 2,5%, 50% a 97,5% percentil, který představuje 95% oboustranný interval spolehlivosti kolem střední varianty prognózy, tzn. kolem 50% percentilu neboli mediánu. Nyní budou popsány výsledky faktorové analýzy, zdůvodněna volba trendu a zpracována dílčí prognóza úmrtnosti mužů a žen.

**Tab. 5: Výsledek faktorové analýzy při hodnocení vztahů parametrů Heligmana-Pollarda u mužů**

Parametr Heligmana- Pollarda	Rotované řešení faktorové analýzy		
	Faktor 1	Faktor 2	Faktor 3
A	0,98	0,03	-0,01
C	0,98	-0,10	-0,05
B	0,94	-0,26	-0,01
H	-0,15	0,98	0,09
F	-0,42	0,61	0,56
G	0,01	-0,93	0,27
E	-0,04	0,12	0,89
D	0,05	-0,28	0,73

**Zdroj:** ČSÚ (2016a), ČSÚ (2015a), ČSÚ (2014a), ČSÚ (2013a), ČSÚ (2012a), ČSÚ (2011a), ČSÚ (2010a), ČSÚ (2009a), ČSÚ (2008a), ČSÚ (2007a), ČSÚ (2006a), ČSÚ (2005a), ČSÚ (2004a), vlastní zpracování.

Tab. 6: Výsledek faktorové analýzy při hodnocení vztahů parametrů Heligmana-Pollarda u žen

Parametr Heligmana-Pollarda	Rotované řešení faktorové analýzy		
	Faktor 1	Faktor 2	Faktor 3
A	0,94	-0,10	0,24
C	0,94	-0,24	-0,06
B	0,94	-0,04	0,16
H	-0,10	0,95	-0,04
F	-0,36	0,73	0,03
G	-0,04	-0,97	0,09
E	0,05	-0,01	0,96
D	0,58	-0,14	0,65

**Zdroj:** ČSÚ (2016b), ČSÚ (2015b), ČSÚ (2014b), ČSÚ (2013b), ČSÚ (2012b), ČSÚ (2011b), ČSÚ (2010b), ČSÚ (2009b), ČSÚ (2008b), ČSÚ (2007b), ČSÚ (2006b), ČSÚ (2005b), ČSÚ (2004b), vlastní zpracování.

Na základě hodnocení vlastních čísel z výstupu SAS bylo zjištěno, že u mužů i žen lze vzájemné vztahy mezi parametry Heligmana-Pollarda popsat třemi faktory. Důležitým hlediskem je také kvalita faktorové analýzy. Aby mohla být faktorová analýza použita, musí vycházet Kaiser-Meyer-Olkinova míra větší než 0,5 (Kaiser, 1974). U mužů vychází tato míra 0,55 a u žen 0,56, tzn. oba modely lze použít, ale jejich kvalita je horší. V tabulce 5 je zobrazeno rotované řešení faktorové analýzy parametrů Heligmana-Pollarda u mužů. Faktor 1 je sycen parametry A, B a C. Faktor 2 zahrnuje parametry F, G a H. Poslední faktor dává do vztahu parametry D a E. V tabulce 6 je výsledek rotovaného řešení faktorové analýzy parametrů Heligmana-Pollarda u žen. Faktor 1, faktor 2 i faktor 3 jsou u žen syceny stejnými parametry jako u mužů.

Tab. 7: Odhad logaritmického trendu faktorových skóre jednotlivých faktorů Heligmana-Pollarda mužů a žen

Faktorové skóre příslušného faktoru	Regresní koeficienty	Odhad	Standardní chyba	T-test	P-hodnota
Faktor 1 muži	$b_0$	-0,85	0,72	-1,17	0,27
	$b_1$	0,49	0,38	1,27	0,23
Faktor 2 muži	$b_0$	-0,98	0,70	-1,40	0,19
	$b_1$	0,57	0,37	1,52	0,16
Faktor 3 muži	$b_0$	1,44	0,61	2,35	0,04
	$b_1$	-0,83	0,33	-2,55	0,03
Faktor 1 ženy	$b_0$	0,04	0,77	0,05	0,96
	$b_1$	-0,02	0,41	-0,05	0,96
Faktor 2 ženy	$b_0$	-1,26	0,65	-1,94	0,08
	$b_1$	0,73	0,35	2,11	0,06
Faktor 3 ženy	$b_0$	0,96	0,71	1,36	0,20
	$b_1$	-0,55	0,37	-1,48	0,17

**Zdroj:** ČSÚ (2016a), ČSÚ (2016b), ČSÚ (2015a), ČSÚ (2015b), ČSÚ (2014a), ČSÚ (2014b), ČSÚ (2013a), ČSÚ (2013b), ČSÚ (2012a), ČSÚ (2012b), ČSÚ (2011a), ČSÚ (2011b), ČSÚ (2010a), ČSÚ (2010b), ČSÚ (2009a), ČSÚ (2009b), ČSÚ (2008a), ČSÚ (2008b), ČSÚ (2007a), ČSÚ (2007b), ČSÚ (2006a), ČSÚ (2006b), ČSÚ (2005a), ČSÚ (2005b), ČSÚ (2004a), ČSÚ (2004b), vlastní zpracování.

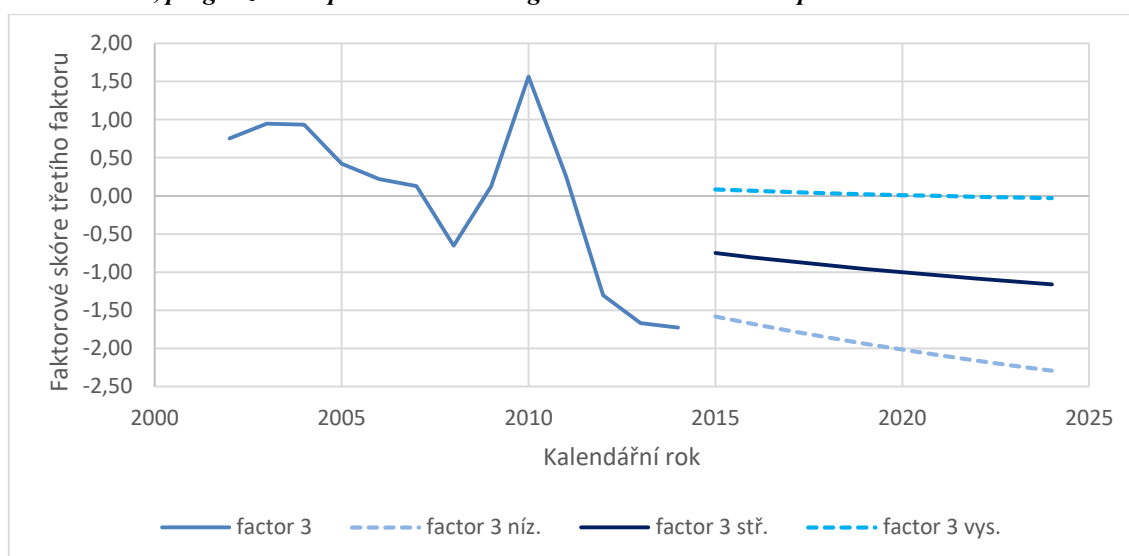
**Pozn.:** Pokud se na 10% hladině spolehlivosti zamítá nulová hypotéza o nevýznamnosti regresního koeficientu  $b_1$  a  $b_0$ , poté je celý řádek označen červenou barvou.

Po volbě vhodného počtu faktorů se vypočítá pro každý faktor jeho faktorové skóre. Jak již bylo předznamenáno, pro odhad očekávaného vývoje faktorového skóre byl zvolen logaritmický

trend, který vychází z povahy dat. V tabulce 7 jsou zaznamenány odhady parametrů logaritmického trendu jednotlivých faktorových skóre pro muže a ženy včetně směrodatné chyby, t-testu a p-hodnoty. Pro hodnocení významnosti parametrů logaritmického trendu byla zvolena hladina spolehlivosti 10 % z důvodu méně kvalitních odhadů způsobených agregací dat. Logaritmický trend je významný u faktorového skóre třetího faktoru u mužů a faktorového skóre druhého faktoru u žen. Ostatní parametry logaritmického trendu faktorových skóre nejsou na 10% hladině spolehlivosti významné, a proto jsou tyto parametry nahrazeny nulou.

Pro přehlednější orientaci v grafech budou v legendě používány následující zkratky: níž. představuje nízkou variantu, stř. symbolizuje střední variantu a vys. značí vysokou variantu. Všechny tyto zkratky jsou také umístěny v seznamu zkratek. Ostatní zkratky budou vždy popsány pod grafem.

**Obr. 12:** Vývoj faktorového skóre třetího faktoru úmrtnosti mužů mezi lety 2002–2014 a jeho prognóza do roku 2024, prognózováno prostřednictvím logaritmického trendu se spolehlivostí 95 %

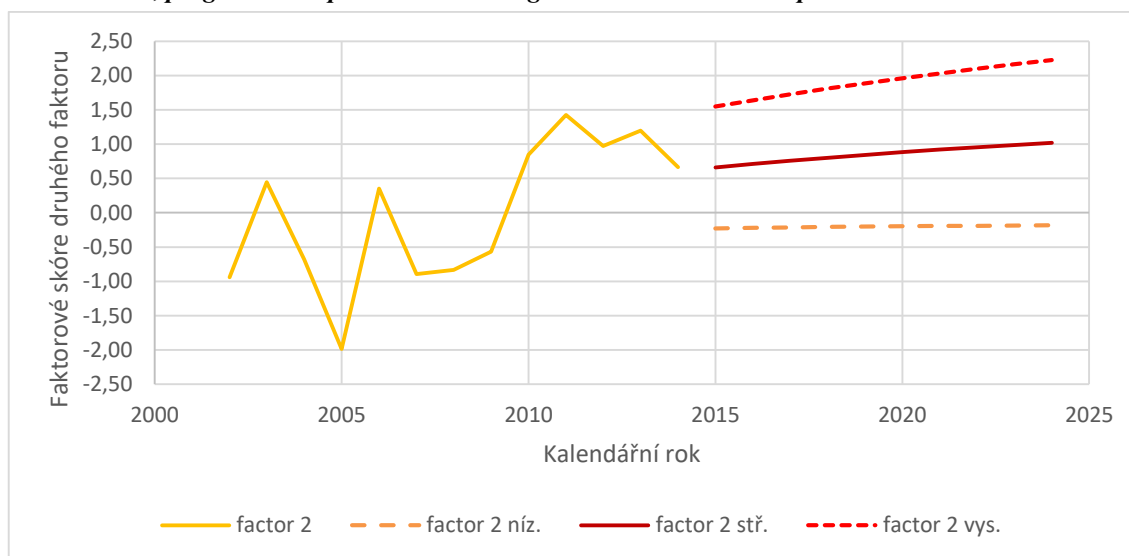


**Zdroj:** ČSÚ (2016a), ČSÚ (2015a), ČSÚ (2014a), ČSÚ (2013a), ČSÚ (2012a), ČSÚ (2011a), ČSÚ (2010a), ČSÚ (2009a), ČSÚ (2008a), ČSÚ (2007a), ČSÚ (2006a), ČSÚ (2005a), ČSÚ (2004a), vlastní zpracování.

Na obrázku 12 je zobrazen vývoj faktorového skóre třetího faktoru u mužů včetně jeho prognózy do roku 2024 s 95% spolehlivostí. Již z obrázku je patrné, že jsou obrovské výkyvy ve vývoji časové řady tohoto faktorového skóre, které jsou způsobené právě nedostatečně kvalitními odhady parametrů Heligmana-Pollarda. Tato skutečnost zapříčiňuje rovněž vychýlenost budoucího odhadu logaritmického trendu tohoto faktorového skóre. Volba logaritmického trendu vychází z povahy dat. Při porovnání vysvětlené variability mezi lineárním trendem a logaritmickým trendem popisuje data lépe logaritmický trend.

Vývoj faktorového skóre druhého faktoru žen včetně jeho prognózy je znázorněn na obrázku 13. I v tomto případě se projevují ne zcela kvalitní odhady parametrů, které mohou být způsobené také rozkolísaností úmrtnostních tabulek z dat Českého statistického úřadu. Pro odhad budoucího vývoje tohoto skóre je zvolen logaritmický trend, který je na 10% hladině spolehlivosti významný. Vzhledem ke skutečnosti, že ostatní faktorová skóre jsou příliš vychýlena špatnými odhady, jsou nahrazena průměrem, který je kvůli z-skóřům nulový.

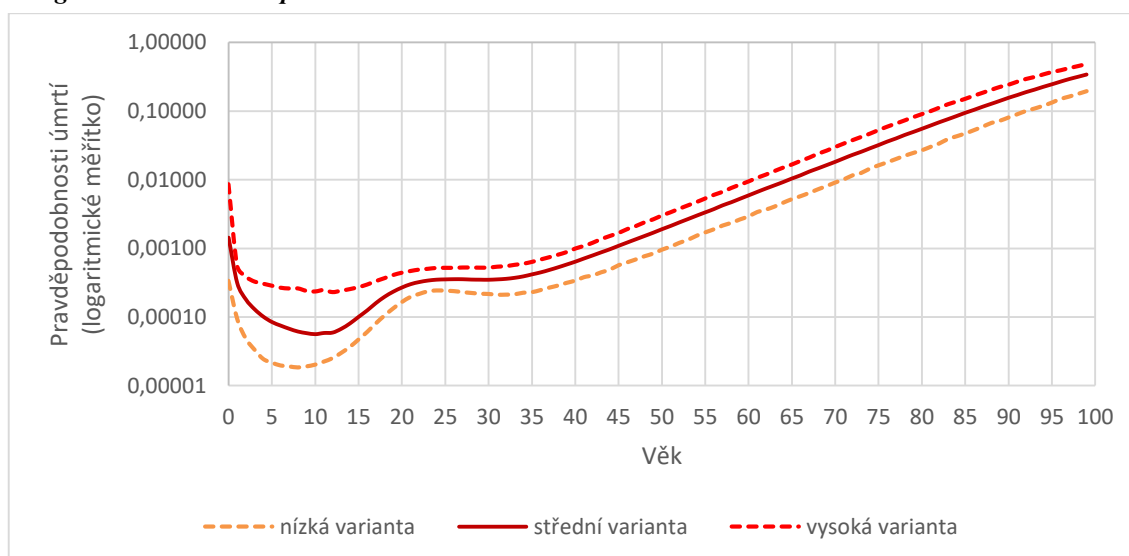
**Obr. 13:** Vývoj faktorového skóre druhého faktoru úmrtnosti žen mezi lety 2002–2014 a jeho prognóza do roku 2024, prognózováno prostřednictvím logaritmického trendu se spolehlivostí 95 %



**Zdroj:** ČSÚ (2016b), ČSÚ (2015b), ČSÚ (2014b), ČSÚ (2013b), ČSÚ (2012b), ČSÚ (2011b), ČSÚ (2010b), ČSÚ (2009b), ČSÚ (2008b), ČSÚ (2007b), ČSÚ (2006b), ČSÚ (2005b), ČSÚ (2004b), vlastní zpracování.

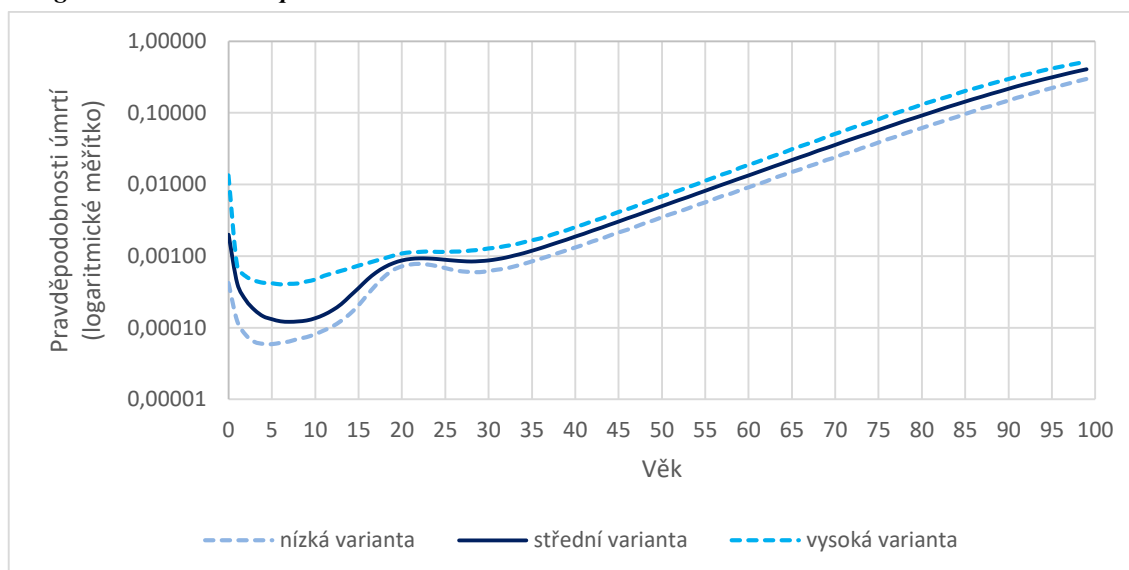
Na obrázku 14 je zpracována prognóza pravděpodobností úmrtí žen v roce 2024 a na obrázku 15 je zobrazena stejná prognóza pro muže. U mužů i u žen se očekává klasický profil úmrtnosti podle věku, kdy nejprve bude vyšší kojenecká úmrtnost. Přibližně ve věkové skupině 5 až 10 let se předpokládá nejnižší úmrtnost. Okolo 21. roku u mužů a 23. roku u žen lze očekávat zvýšení úmrtnosti z důvodu vnějších příčin. Přibližně od věku 35 let u žen a 30 let u mužů lze předpokládat lineární růst úmrtnosti. Lineární růst je způsobený logaritmickou transformací osy grafu.

**Obr. 14:** Očekávaný věkový profil pravděpodobností úmrtí žen v roce 2024 zpracovaný modelem Heligmana-Pollarda se spolehlivostí 95 %



**Zdroj:** ČSÚ (2016a), ČSÚ (2015a), ČSÚ (2014a), ČSÚ (2013a), ČSÚ (2012a), ČSÚ (2011a), ČSÚ (2010a), ČSÚ (2009a), ČSÚ (2008a), ČSÚ (2007a), ČSÚ (2006a), ČSÚ (2005a), ČSÚ (2004a), vlastní zpracování.

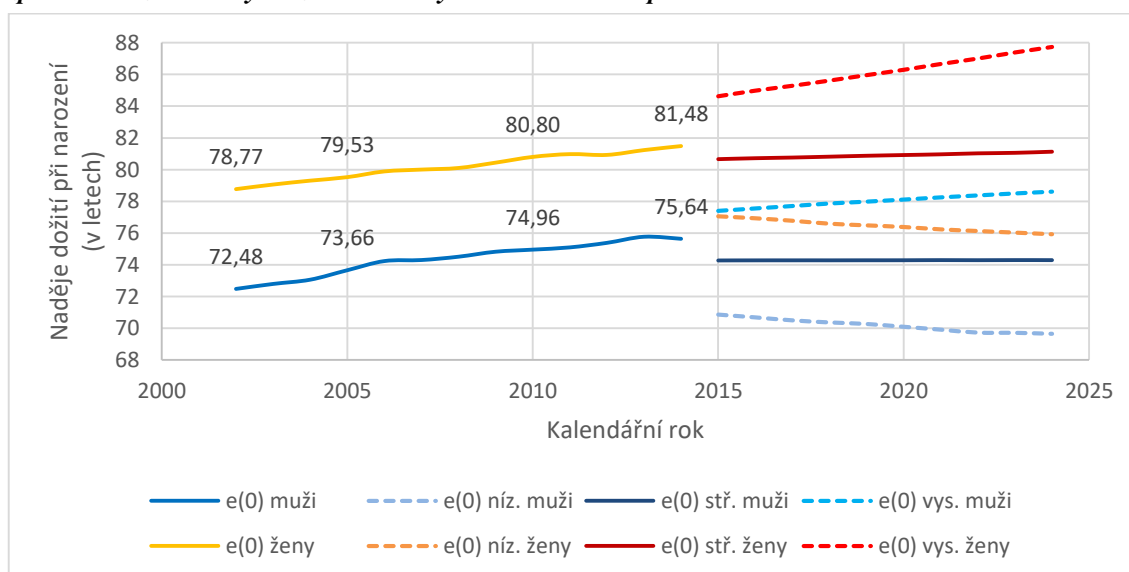
**Obr. 15: Očekávaný věkový profil pravděpodobností úmrtí mužů v roce 2024 zpracovaný modelem Heligmana-Pollarda se spolehlivostí 95 %**



**Zdroj:** ČSÚ (2016a), ČSÚ (2015a), ČSÚ (2014a), ČSÚ (2013a), ČSÚ (2012a), ČSÚ (2011a), ČSÚ (2010a), ČSÚ (2009a), ČSÚ (2008a), ČSÚ (2007a), ČSÚ (2006a), ČSÚ (2005a), ČSÚ (2004a), vlastní zpracování.

Kolem střední varianty prognózy jsou jak u žen (Obr. 14), tak i u mužů (Obr. 15) zobrazeny intervaly spolehlivosti kolem střední hodnoty s 95% spolehlivostí. Velká šířka těchto intervalů je způsobena kvůli nedokonalosti odhadů parametrů Heligmana-Pollarda. Tato skutečnost může být inspirací pro další práce k vytvoření programu pro odhad parametrů Heligmana-Pollarda z jednoletých věkových intervalů.

**Obr. 16: Vývoj naděje dožití chlapců a dívek při narození mezi lety 2002–2014 a jejich prognóza zpracovaná z krátké výchozí časové řady do roku 2024 se spolehlivostí 95 %**



**Zdroj:** ČSÚ (2016a), ČSÚ (2016b), ČSÚ (2015a), ČSÚ (2015b), ČSÚ (2014a), ČSÚ (2014b), ČSÚ (2013a), ČSÚ (2013b), ČSÚ (2012a), ČSÚ (2012b), ČSÚ (2011a), ČSÚ (2011b), ČSÚ (2010a), ČSÚ (2010b), ČSÚ (2009a), ČSÚ (2009b), ČSÚ (2008a), ČSÚ (2008b), ČSÚ (2007a), ČSÚ (2007b), ČSÚ (2006a), ČSÚ (2006b), ČSÚ (2005a), ČSÚ (2005b), ČSÚ (2004a), ČSÚ (2004b), vlastní zpracování.

Na obrázku 16 je zobrazena časová řada skutečné naděje dožití při narození chlapců a dívek, včetně prognózy zpracované z krátké výchozí časové řady s 95% spolehlivostí. Do naděje dožití se promítají odhady parametrů Heligmana-Pollarda s omezenou kvalitou, které způsobují vychýlenost střední varianty prognózy a velkou šíři intervalů spolehlivosti. Z tohoto důvodu zde nebude více přiblížen očekávaný demografický vývoj, který je méně spolehlivý, ale bude popsáno na střední variantě prognózy kvalitní metodické zpracování této problematiky. Střední varianta prognózy naděje dožití mužů i žen je podhodnocena oproti skutečné naději dožití v posledním známém roce. Tato skutečnost je způsobena tím, že faktorové skóre prvního faktoru se prognózuje konstantním trendem a to způsobuje jeho výrazné odchýlení od skutečného vývoje. Toto vychýlení se kvůli vztahu mezi prvním faktorem a parametry A, B a C Heligmana-Pollarda tzn. parametry dětské úmrtnosti přenáší právě na dětskou úmrtnost. Protože má dětská úmrtnost nejvyšší vliv na naději dožití při narození, dochází v tomto případě k výraznému snížení naděje dožití při narození. U žen byla statisticky prokázána snižující se tendence úmrtnosti v nejvyšších věcích a rostoucí věk lokalizace úmrtnosti na vnější příčiny, tzn. byl prokázán logaritmický trend u faktorového skóre druhého faktoru. Prostřednictvím vztahu mezi faktorovým skóre druhého faktoru a parametry F, G a H Heligmana-Pollarda se toto snížení úmrtnosti a zvýšení věku lokalizace úmrtnosti na vnější příčiny projevuje v mírném nárůstu střední varianty prognózy naděje dožití při narození u žen. U mužů se tato tendence statisticky neprokázala a dochází rovněž k predikci lineárním trendem u faktorového skóre druhého faktoru. Jako statistiky významný se u mužů jeví logaritmický trend u faktorového skóre třetího faktoru. Toto faktorové skóre má vliv na parametry D a E Heligmana-Pollarda. Kvůli výkyvu faktorového skóre třetího faktoru v roce 2010 (Obr. 12) je odhad tohoto skóre velmi nadhodnocen a dochází pouze k jeho mírnému poklesu. Tento mírný pokles prakticky nemá vliv na naději dožití při narození mužů. U žen se pokles faktorového skóre druhého faktoru neprokázal.

## 6.2 Dílčí prognóza plodnosti s využitím Beta funkce

Pro dílčí prognózu měr plodnosti, která vychází z krátké výchozí časové řady, byla zvolena analýza prostřednictvím Beta funkce. Parametry této funkce lze následně prognózovat. Problémem ale je, že mezi parametry Beta funkce mohou být vzájemné vztahy, a proto nelze modelovat každý parametr zvlášť. Stejným způsobem jako u dílčí prognózy úmrtnosti byla k identifikaci vazeb mezi parametry použita faktorová analýza.

**Tab. 8: Vlastní čísla z faktorové analýzy při hodnocení počtu nezávislých parametrů Beta funkce**

Vlastní číslo	Rozdíl	Poměr	Kumulace
2,81	2,67	0,94	0,94
0,13	0,07	0,04	0,98
0,06	–	0,02	1,00

**Zdroj:** ČSÚ (2016c), ČSÚ (2010c), ČSÚ (2008c), SAS Institute Inc (2017).

Při hodnocení vlastních čísel z faktorové analýzy vyplývá (Tab. 8), že mezi všemi parametry Beta funkce existují významné vazby a tyto vazby lze podchytit jedním faktorem. Tuto skutečnost



lze vysledovat z vlastních čísel, kdy jejich hodnoty větší než jedna tvoří samostatný faktor. Podle Kaiser-Meyer-Olkinovy míry, která vychází pro tento model 0,77, lze faktorovou analýzu použít. Následně je vypočítáno faktorové skóre pro tento jeden faktor, které nahrazuje hodnoty původních tří proměnných a vytváří nové prostřednictvím vzájemných vazeb. Poté je nutné vyčíslit vztah mezi faktorovým skóre a původními parametry Beta funkce upravenými na z-skóry. K tomuto účelu je využita jednoduchá lineární regrese, kde za  $x$  se dosadí faktorová skóre a za  $y$  se postupně dosadí upravené parametry Beta funkce. Výsledkem je nulový průsečík s osou  $y$  a sklon je totožný s rotovaným řešením faktorové analýzy. Vzhledem ke skutečnosti, že pro jeden faktor není možné provést rotaci, pak je tímto sklonem nerotované řešení faktorové analýzy (Tab. 9).

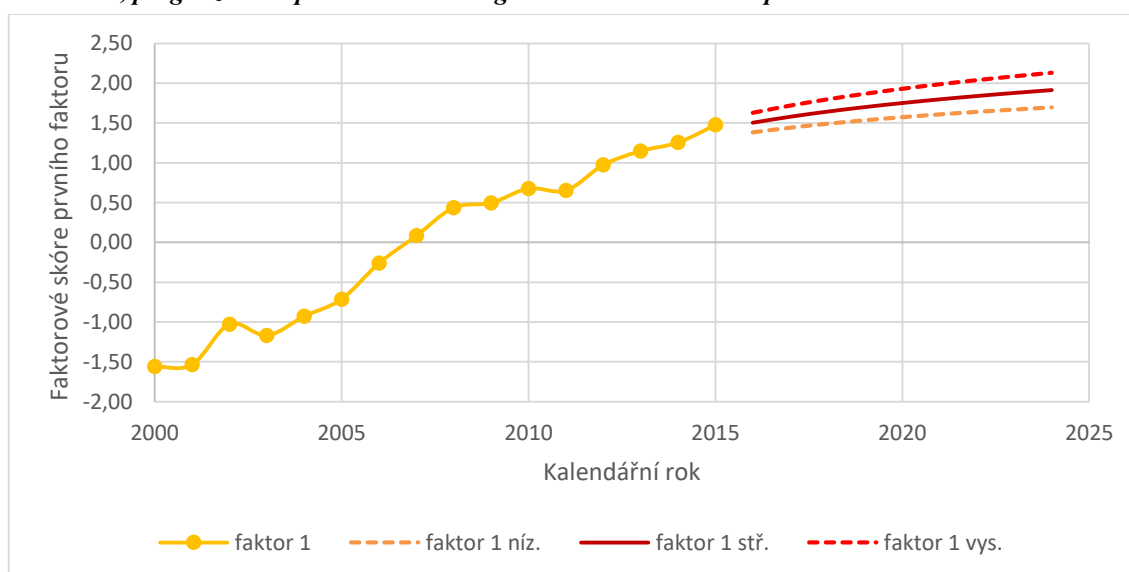
**Tab. 9: Výsledek faktorové analýzy při hodnocení vztahů parametrů Beta funkce**

Beta funkce	Faktor 1
Průměrný věk	0,98
Úhrnná plodnost	0,97
Rozptyl	0,96

**Zdroj:** ČSÚ (2016c), ČSÚ (2010c), ČSÚ (2008c), vlastní zpracování.

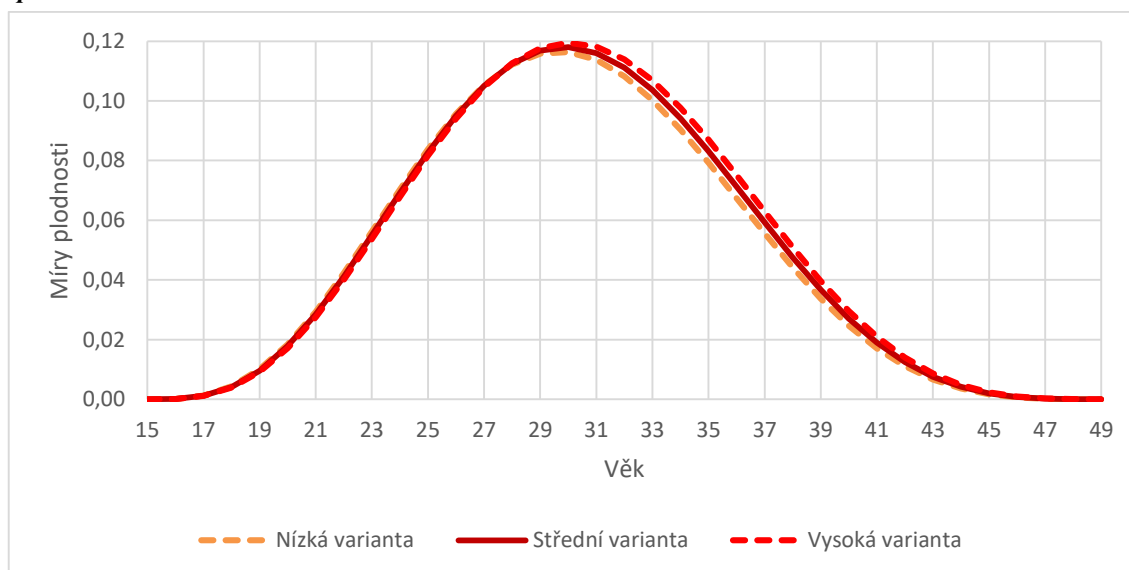
Na obrázku 17 je zobrazen vývoj faktorového skóre u plodnosti mezi roky 2000 až 2015. Od roku 2011 vykazuje přibližně logaritmický trend a z tohoto důvodu je vytvořena prognóza faktorového skóre od roku 2016 do roku 2024 právě prostřednictvím logaritmického trendu včetně 95% intervalů spolehlivosti. Z odhadnutých faktorových skóre jsou podle výše popsané jednoduché lineární regrese pro všechny tři varianty prognózy zpětně odhadnuty jednotlivé parametry Beta funkce. Je potřeba si uvědomit, že tato data jsou upravená na z-skóry, a proto je nutné vynásobit je příslušnou směrodatnou odchylkou a přičíst jejich průměr.

**Obr. 17: Vývoj faktorového skóre prvního faktoru u plodnosti mezi lety 2000–2015 a jeho prognóza do roku 2024, prognózováno prostřednictvím logaritmického trendu se spolehlivostí 95 %**



**Zdroj:** ČSÚ (2016c), ČSÚ (2010c), ČSÚ (2008c), vlastní zpracování.

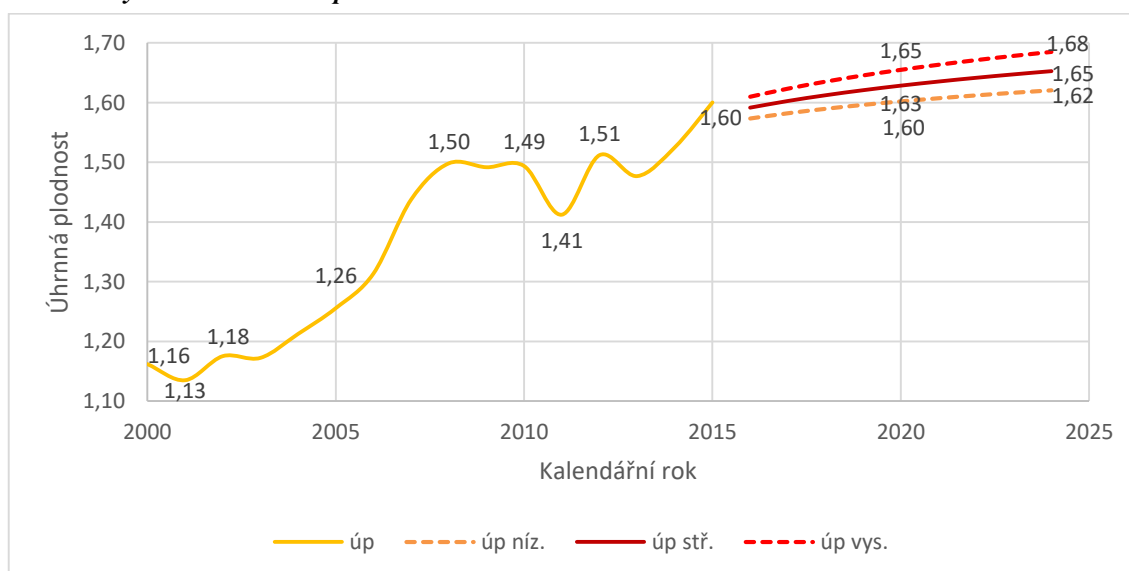
**Obr. 18:** Očekávané míry plodnosti podle věku v roce 2024 zpracované prostřednictvím Beta funkce se spolehlivostí 95 %



**Zdroj:** ČSÚ (2016c), ČSÚ (2010c), ČSÚ (2008c), vlastní zpracování.

Na obrázku 18 je zobrazena prognóza měr plodnosti v roce 2024 v Jihočeském kraji zpracovaná z krátké výchozí časové řady. Podle střední varianty prognózy lze očekávat průměrný věk žen při narození dítěte 30,4 let. Vysoká varianta předpokládá průměrný věk při narození dítěte na úrovni 30,6 let a nízká varianta očekává, že tento věk bude přibližně 30,2 let. Na základě tohoto očekávání dochází k mírnému posunu křivky plodnosti v rámci variant.

**Obr. 19:** Vývoj úhrnné plodnosti mezi lety 2000–2015 a její prognóza zpracovaná z krátké výchozí časové řady do roku 2024 se spolehlivostí 95 %



**Zdroj:** ČSÚ (2016c), ČSÚ (2010c), ČSÚ (2008c), vlastní zpracování.

**Pozn.:** úp představuje úhrnnou plodnost.

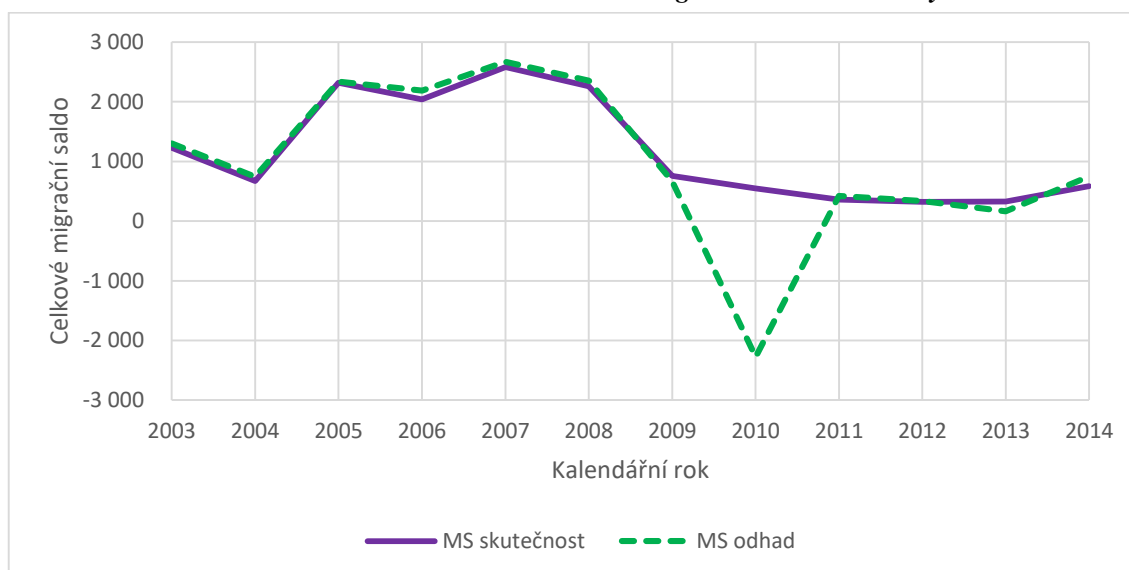
Na obrázku 19 je zobrazena prognóza vývoje úhrnné plodnosti v Jihočeském kraji do roku 2024 včetně původní časové řady tohoto ukazatele. Nejvyšší úroveň úhrnné plodnosti byla

v období 2000 až 2015 zaznamenána v roce 2001, naopak nejvyšší úroveň byla zjištěna v roce 2015, a to v průměru 1,60 dětí na jednu ženu. Do budoucna lze podle tohoto vývoje očekávat nárůst úhrnné plodnosti. Podle této prognózy by úhrnná plodnost měla být v roce 2024 mezi 1,62 a 1,68 dětí na jednu ženu.

### 6.3 Výpočet migračního salda prostřednictvím nepřímé metody včetně prognózy jeho očekávaného vývoje

Vzhledem ke skutečnosti, že z internetových stránek ČSÚ nejsou k dispozici data za migraci Jihočeského kraje podle jednotek věku, bylo nutné tato data odhadnout. Pro odhad byla zvolena nepřímá metoda výpočtu migračního salda. Alternativní možností byl výpočet migračního salda z dat České republiky prostřednictvím počtu přistěhovalých a vystěhovalých podle jednotek věku za předpokladu, že věková struktura migračního salda v Jihočeském kraji je shodná s věkovou strukturou migračního salda České republiky. Počty přistěhovalých a vystěhovalých celé České republiky jsou k dispozici ze stránek ČSÚ. K této alternativní možnosti nakonec nebylo přistoupeno z důvodu, že do Jihočeského kraje mohou přicházet lidé i z ostatních krajů, nejen ze zahraničí a jejich struktura se může lišit.

**Obr. 20:** Porovnání skutečného a odhadnutého celkového migračního salda mezi lety 2003–2014



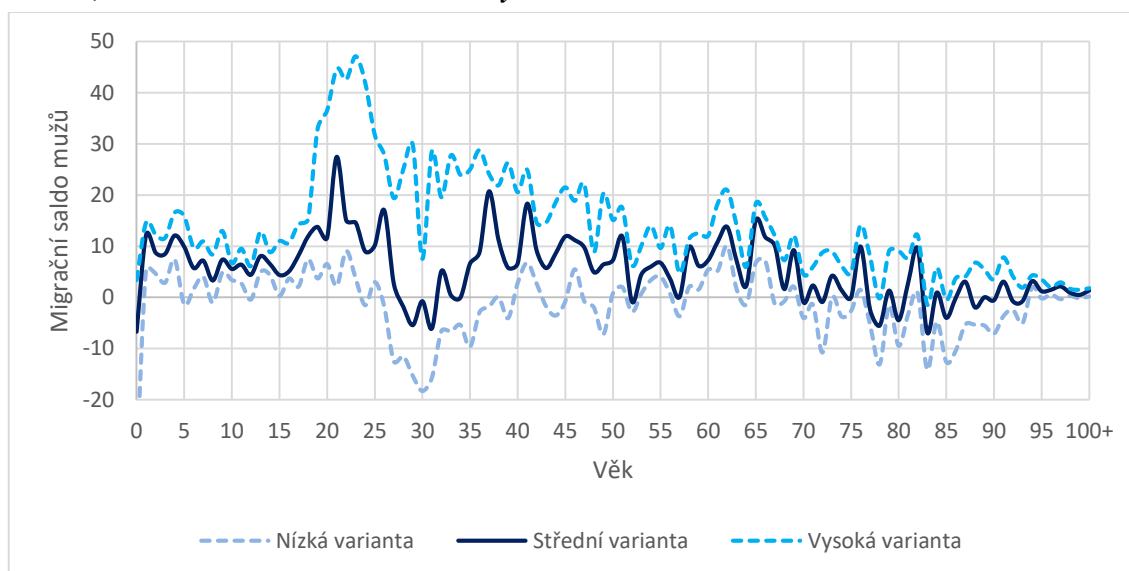
**Zdroj:** ČSÚ (2016f), vlastní zpracování.

**Pozn.:** MS představuje celkové migrační saldo.

Po odhadu migračního salda mužů a žen podle jednotek věku nepřímou metodou výpočtu (Kap. 3.4) bylo z těchto dat zpracováno celkové migrační saldo obyvatel Jihočeského kraje. Porovnáním takto odhadnutého celkového migračního salda se skutečnými hodnotami tohoto celkového salda z dat ČSÚ (Obr. 20), byl zjištěn výrazný výkyv v roce 2010. Rozdíly mezi skutečnými a odhadnutými hodnotami jsou u ostatních roků malé. Po konzultaci s doktorkou Pavlínou Habartovou z ČSÚ bylo zjištěno, že stav žijících mužů a žen v Jihočeském kraji byl

upraven k 1. 1. 2011 z důvodu zpřesnění informací ze Sčítání lidu, domů a bytů z roku 2011. Všechny změny jsou uveřejněny na stránkách ČSÚ (2017b). Porovnáním odhadnutého migračního salda k 1. 1. 2011 vypočteného nepřímou metodou a skutečného migračního salda k 1. 1. 2011 upraveného o sčítání lidu, dochází k výrazné změně migračního salda v roce 2010 způsobené nikoli migrací, ale zpřesněním prostřednictvím sčítání. Z tohoto důvodu bylo odhadnuté migrační saldo z roku 2010 označeno jako odlehlé pozorování a bylo odstraněno ze souboru. Následně byla odhadnutá časová řada migračního salda podle jednotek věku analyzována prostřednictvím dolního kvartilu, mediánu a horního kvartilu, které budou promítnuty do populační prognózy Jihočeského kraje. Podrobnosti o technickém zpracování prognózy migračního salda byly popsány v teoretické části práce.

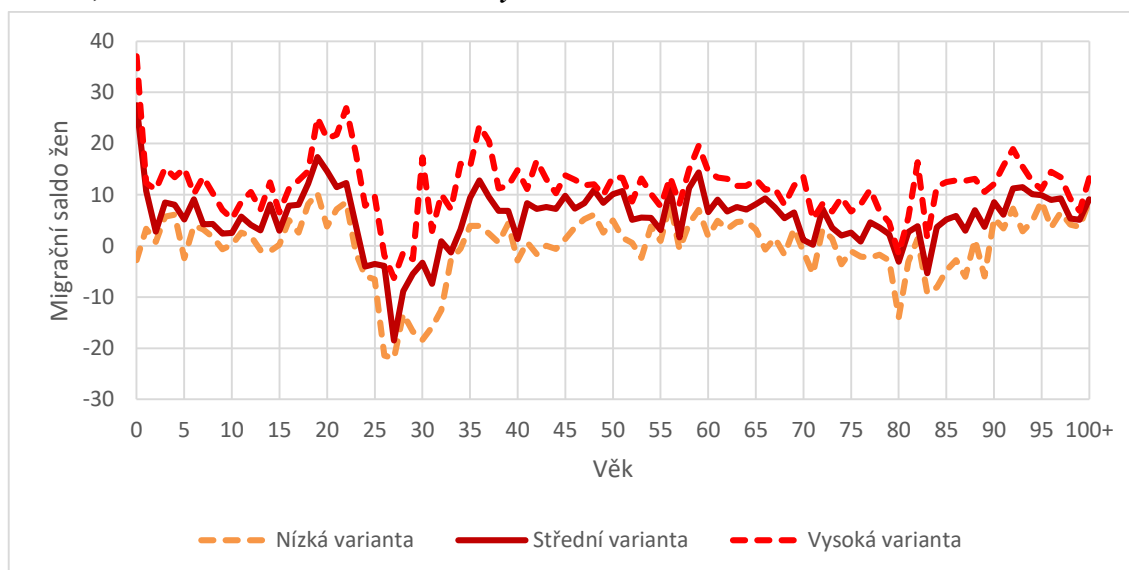
**Obr. 21:** *Odhad migračního salda mužů podle jednotek věku, prognózováno prostřednictvím dolního kvartilu, mediánu a horního kvartilu mezi lety 2015–2024*



**Zdroj:** vlastní zpracování.

Na obrázku 21 je zpracována prognóza struktury migračního salda mužů v Jihočeském kraji, která se očekává mezi roky 2015 až 2020. Stejným způsobem je zpracována prognóza struktury migračního salda i u žen a její výsledky jsou zobrazeny na obrázku 22. Všeobecným předpokladem je, že struktura migračního salda se bude pohybovat právě mezi dolním a horním kvantilem. Pro upřesnění dolní kvartil představuje v prognóze struktury migračního salda nízkou variantu prognózy, medián symbolizuje nejpravděpodobnější střední variantu a horní kvartil značí vysokou variantu. Nejvyšší přírůstek podle střední varianty migračního salda mužů a žen se očekává ve věku 20 let. Může to být způsobeno dočasným nárůstem počtu studentů. Naopak výrazný pokles migračního salda lze předpokládat přibližně ve věkové skupině 25 až 30 let. Tato skutečnost může být zapříčiněna odchodem studentů za prací mimo Jižní Čechy. Rozdíl migračního salda mezi chlapci a dívkami při narození může být způsobený například jiným podílem děvčat při narození kvůli menšímu území Jihočeského kraje. Druhou možností tohoto rozdílu může být nezapočítání dětí narozených přistěhovalým ženám. V ostatních věcech se očekává, že u migračního salda mužů a žen bude docházet k relativně rovnoměrným změnám.

**Obr. 22: Odhad migračního salda žen podle jednotek věku, prognózováno prostřednictvím dolního kvartilu, mediánu a horního kvartilu mezi lety 2015–2024**



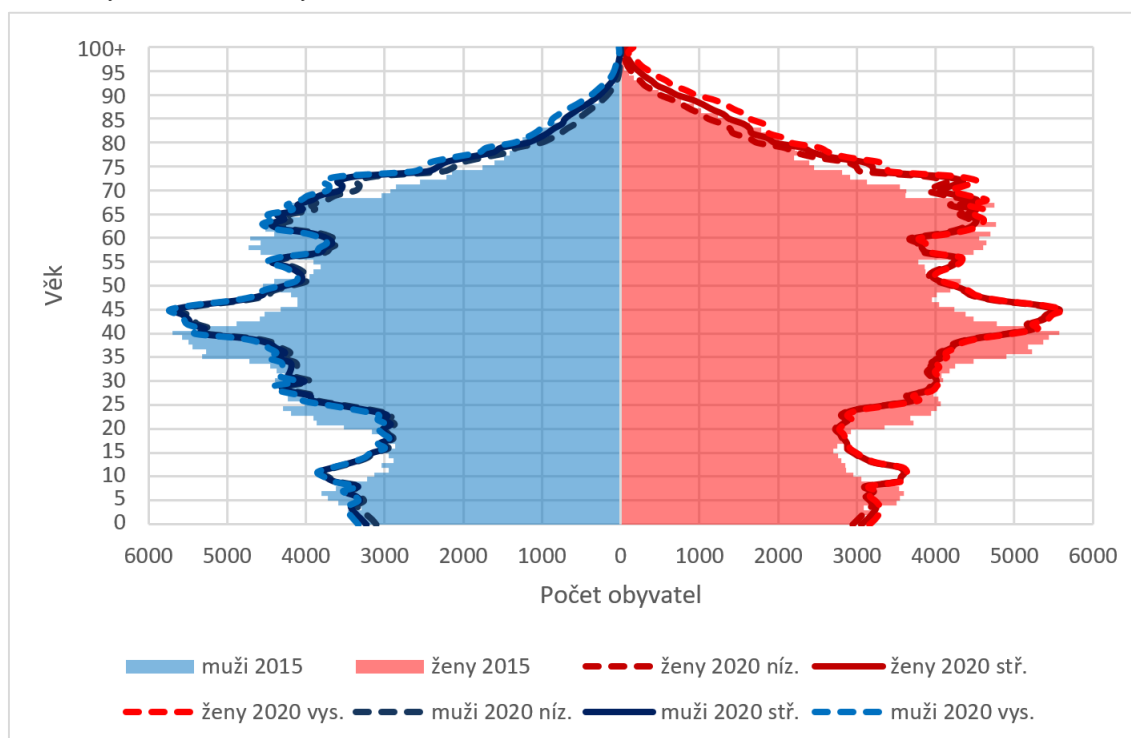
**Zdroj:** vlastní zpracování.

#### **6.4 Populační prognóza počtu a věkové struktury mužů a žen v Jihočeském kraji zpracovaná z dílčích prognóz úmrtnosti, plodnosti a migrace analýzou krátké výchozí časové řady**

Doposud byla popsána dílčí prognóza úmrtnosti, plodnosti a migrace z krátké výchozí časové řady. Závěrečnou etapou je vytvoření populační prognózy Jihočeského kraje z těchto dílčích prognóz. Tato prognóza je zpracována kohortně komponentní metodou s migrací s dvoufázovým zahrnutím migračního salda. Samozřejmě je nutné nezapomenout na to, že prognóza z krátké výchozí časové řady má pouze malou vypovídající schopnost z důvodu méně kvalitních vstupujících dat. Dalším problémem je velmi krátká výchozí časová řada, ve které se velmi obtížně analyzují jakékoli trendy. Dále do prognózy vstupují někdy ne zcela kvalitní odhady parametrů dílčích prognóz, které se následně promítají do celkové populační prognózy. Vzhledem k těmto překážkám nebude populační prognóza Jihočeského kraje vytvořená z krátké výchozí časové řady porovnána v osmé kapitole s populační prognózou vytvořenou Českých statistickým úřadem. Přesto použitím kvalitních modelů v dílčích prognózách lze řadu z těchto problémů vyřešit. Přesnější populační prognóza Jihočeského kraje bude zpracována v kapitole 7.

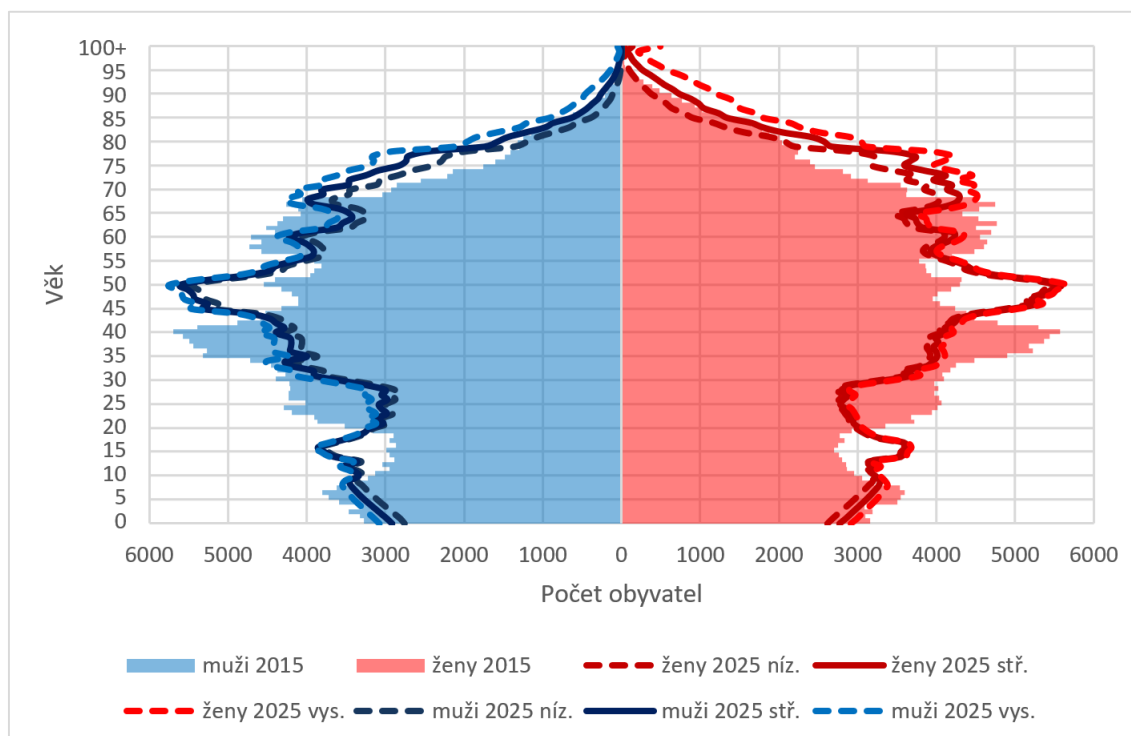
Na obrázku 23 je zobrazena prognóza věkové struktury, která se očekává k 1. 1. 2020 a na obrázku 24 je zpracována tato předpokládaná věková struktura o pět let později v porovnání s rokem 2015. Na věkové struktuře k 1. 1. 2015 bylo nejvíce obyvatel ve věku 40 let. Jedná se o generace narozené v 70. letech 20. století, kdy se prostřednictvím výrazných propopulačních opatření a silné generaci žen v produktivním věku rodilo více dětí. Po roce 2000 docházelo také k mírnějšímu nárůstu počtu narozených, které bylo způsobené odkladem mateřství do vyššího věku. Podrobněji tuto problematiku popisuje Kocourková (2008, s. 240–249).

**Obr. 23:** Porovnání prognózy věkové struktury mužů a žen k 1. 1. 2020 oproti 1. 1. 2015 zpracované z krátké výchozí časové řady



**Zdroj:** ČSÚ (2015c), ČSÚ (2015d), vlastní zpracování.

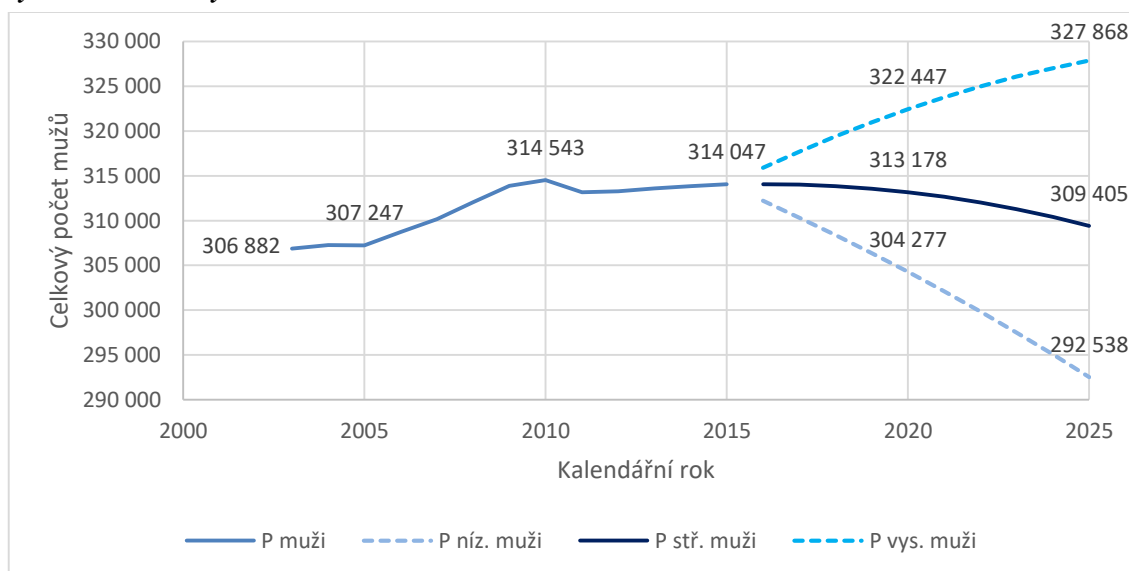
**Obr. 24:** Porovnání prognózy věkové struktury mužů a žen k 1. 1. 2025 oproti 1. 1. 2015 zpracované z krátké výchozí časové řady



**Zdroj:** ČSÚ (2015c), ČSÚ (2015d), vlastní zpracování.

V roce 2020 (Obr. 23) lze očekávat postupné stárnutí věkové struktury a snižování počtu živě narozených dětí. V rámci variant jsou nejvyšší rozdíly v nejvyšších věcích, kdy lze očekávat různé úmrtnostní vzorce. Nízká varianta očekává vysokou intenzitu úmrtnosti a vysoká varianta předpokládá celkové zvýšení počtu obyvatel, a tedy nízkou úmrtnost. V nižších věcích není změna úmrtnosti v rámci variant tak výrazná na věkovou strukturu, protože je intenzita úmrtnosti v těchto věcích nižší. Další výraznější změna v rámci variant se předpokládá v nejnižších věcích, kdy se očekává různá intenzita plodnosti. V roce 2025 (Obr. 24) se oproti roku 2020 ještě prohloubí stárnutí obyvatel. Rovněž narůstá nejistota při dílčí prognóze úmrtnosti, plodnosti a migrace v tomto roce, která způsobuje výraznější nejistotu v rámci jednotlivých variant prognózy. Očekává se také, že počty narozených budou výrazně klesat.

**Obr. 25:** Vývoj celkového počtu mužů k 1. 1. mezi lety 2003–2015 a jeho prognóza zpracovaná z krátké výchozí časové řady do roku 2025

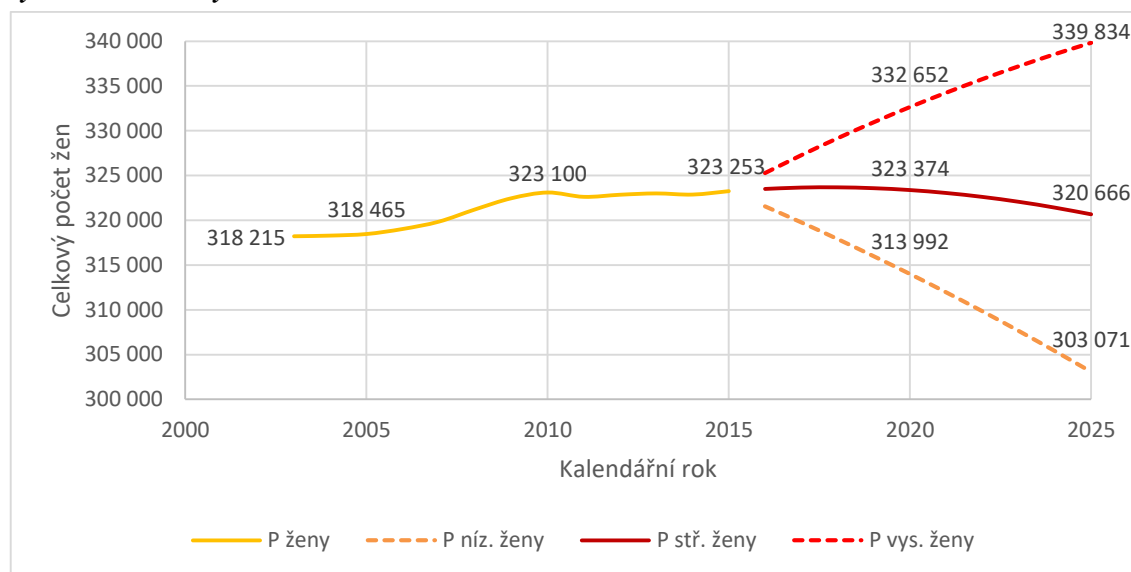


**Zdroj:** ČSÚ (2016d), ČSÚ (2015c), ČSÚ (2014c), ČSÚ, (2013c), ČSÚ (2012c), ČSÚ (2011c), ČSÚ (2010d), ČSÚ (2009c), ČSÚ (2008d), ČSÚ, (2007c), ČSÚ (2006c), ČSÚ (2005c), ČSÚ (2004c), vlastní zpracování.

**Pozn.:** P představuje celkový počet mužů k 1. 1.

Na obrázku 25 je zobrazena prognóza celkového počtu mužů v Jihočeském kraji z krátké výchozí časové řady. Na obrázku 26 je zpracována stejná prognóza pro ženy. Z minulého průběhu časové řady je patrné, že se mezi lety 2003 až 2015 celkový počet osob v Jihočeském kraji zvýšil. Podle střední varianty prognózy se očekává u mužů i u žen výrazný pokles jejich počtu do roku 2025. Tato skutečnost může být především u mužů způsobena méně kvalitními odhady parametrů Heligmana-Pollarda, kdy při hledání zákonitostí nebyla na 10% hladině spolehlivosti potvrzena snižující se tendence úmrtnosti v nejvyšších věcích, jako tomu je u žen. Tato skutečnost poté může podhodnocovat celkový počet mužů ve střední variantě. Nízká i vysoká varianta jsou málo pravděpodobné. Z minulého průběhu časové řady vyplývá, že je celkově více žen v populaci než mužů. To lze přisuzovat mužské nadúmrtosti, která je promítnuta i do prognózy celkového počtu mužů a žen.

**Obr. 26:** Vývoj celkového počtu žen k 1. 1. mezi lety 2003–2015 a jeho prognóza zpracovaná z krátké výchozí časové řady do roku 2025



**Zdroj:** ČSÚ (2016e), ČSÚ (2015d), ČSÚ (2014d), ČSÚ (2013d), ČSÚ (2012d), ČSÚ (2011d), ČSÚ (2010e), ČSÚ (2009d), ČSÚ (2008d), ČSÚ (2007c), ČSÚ (2006c), ČSÚ (2005c), ČSÚ (2004c), vlastní zpracování.

**Pozn.:** P představuje celkový počet mužů k 1. 1.



## Kapitola 7

### Prognóza vývoje počtu obyvatel Jihočeského kraje prostřednictvím analyzování vývoje dlouhé výchozí časové řady

Cílem této kapitoly je vytvoření populační prognózy Jihočeského kraje prostřednictvím analyzování vývoje dlouhé výchozí časové řady. Prostředkem pro zpracování této prognózy budou dílčí prognózy úmrtnosti, plodnosti a migrace, kde migrace bude analyzována zvlášť za přistěhované a zvlášť za vystěhované jedince. Pro tvorbu těchto dílčích prognóz byl zvolen Lee-Carterův model, který se skládá z indexu závislého na čase a dvou parametrů závislých na věku. Věkové parametry jsou v čase konstantní a jediným parametrem závislým na čase je index úrovně úmrtnosti, plodnosti, imigrace a emigrace dále jen index (Habartová, 2009, s. 174). Tento index se modeluje prostřednictvím analýzy časových řad, kde výběr metody závisí na délce výchozí časové řady. Pro plodnost a úmrtnost byla zvolena sofistikovanější Box-Jenkinsova metodologie, kde předpokladem užití je minimální délka časové řady o 30 pozorování a tuto podmínku výchozí časová řada splňuje. Pro migraci je k dispozici výchozí časová řada o 23 pozorování. V tomto případě již nelze užít Box-Jenkinsovu metodologii, ale pro prognózu indexu byla zvolena analýza prostřednictvím trendových funkcí. K sestavení výsledné populační prognózy byla zvolena kohortně komponentní metoda s migrací s dvoufázovým zahrnutím imigrace a emigrace do modelu.

Dílčí prognózy jsou zpracovány v programu R (R Core Team, 2016), do kterého se doinstalují balíčky Demography (Booth a kol., 2014) a Forecast (Hyndman, 2016). První balíček slouží pro odhad parametrů Lee-Carterova modelu, kdy prostřednictvím funkce (Booth a kol., 2017, s. 21)

$$LCA = lca(DATA, interpolate = TRUE) \quad (7.1)$$

se odhadnou parametry Lee-Carterova modelu. V příkazu (7.1) symbolizuje *lca* Lee-Carterův model, *DATA* značí specifický soubor demografických dat, *interpolate=TRUE* představuje vyrovnaní nulových dat interpolací. Specifický soubor demografických dat se vytvoří příkazem (Booth a kol., 2017, s. 8–9)

$$DATA = demogdata(data = RATE, pop = POP, ages = AGES, years = YEARS, type = "mortality", label = "JC", name = "Male"). \quad (7.2)$$

V instrukci (7.2) představují *demogdata* soubor demografických dat, kde *RATE* je matice měr (v tomto případě tabulkových měr úmrtnosti mužů), *POP* je matice středních stavů obyvatel (v tomto případě se jedná o matici středních stavů mužů), *AGES* je vektor věků, *YEARS* je vektor kalendářních roků, *mortality* značí úmrtnost, *JC* je popisek Jižních Čech a *Male* označuje muže.

Druhý balíček s názvem *Forecast* slouží pro analýzu časových řad prostřednictvím Box-Jenkinsovy metodologie. V této práci bude tato metodologie využita pro prognózu indexu Lee-Carterova modelu. Prostřednictvím příkazu (Hyndman, 2017, s. 11)

*forecast(auto.arima(LCA\$kt), h = 20, level = c(95))* (7.3)

se nejjednodušeji získají odhady indexu s určitou spolehlivostí. Princip příkazu (7.3) spočívá v tom, že prostřednictvím *forecast* se odhadnou konkrétní hodnoty indexu (v tomto případě indexu úrovně úmrtnosti), pomocí příkazu *auto.arima* se bez hlubších znalostí statistických metod vybere na základě určitého algoritmu nejlepší řád  $p$ ,  $d$  a  $q$  smíšeného ARIMA( $p,d,q$ ) modelu, podle kterého bude časová řada indexu (v tomto případě indexu úrovně úmrtnosti) modelována,  $h$  značí časové období, pro které bude prognóza zpracována (v tomto případě na 20 let) a *level* určuje spolehlivost prognózy (v tomto případě 95 %). Celková prognóza je na závěr vytvořena v programu Excel.

## 7.1 Dílčí prognóza úmrtnosti zpracovaná s využitím Lee-Carterova modelu úmrtnosti

Prognóza úmrtnosti je zpracována zvlášť pro muže a zvlášť pro ženy prostřednictvím Lee-Carterova modelu úmrtnosti, kde index úrovně úmrtnosti je modelován pomocí ARIMA modelu. Popsána zde bude dílčí prognóza úmrtnosti Jihočeského kraje s 95% spolehlivostí. Pro praktické účely je ale vhodnější prognóza úmrtnosti Jihočeského kraje s 75% spolehlivostí, která je zobrazena v přílohách 4, 5 a 6.

K vytvoření úmrtnostních tabulek z dat Českého statistického úřadu byl využit program DeRaS (Burcin a kol., 2012). Prostřednictvím tohoto programu jsou vyrovnány úmrtnostní tabulky v nejvyšších věcích prostřednictvím celé řady funkcí, mezi které patří Coale-Kischer, Gompertz, Gompertz-Makeham, Heligman-Pollard, Kannisto a Thatcher. Všechny funkce jsou extrapolovány do věku 120 let. Pro účely této práce budou použity vyrovnané kvocienty úmrtnosti pouze do věku 100 let. Pomocí Akaikeho informačního kritéria známého pod zkratkou AIC se jako nejvhodnější vybere taková funkce, která má hodnotu tohoto kritéria nejnížší (Pecáková, 2007, s. 90).

Vzhledem k tomu, že Lee-Carterův model sleduje věkově specifické změny v čase a volbou optimální vyrovnávací funkce v každém roce by docházelo ke značnému kolísání způsobeného přechody mezi jednotlivými funkcemi, je proto optimální vybrat pouze jednu vyrovnávací funkci pro všechny roky. Zvolena je vždy taková funkce, která je podle AIC nejčastěji vybrána jako optimální. Následně se vypočítá odhad parametrů Lee-Carterova modelu úmrtnosti prostřednictvím příkazu (7.1) a vypočítá se prognóza indexu úmrtnosti pomocí volby (7.3).

**Obr. 27: Akaikeho informační kritérium při volbě vyrovnávací funkce u mužů mezi lety 1986–2015**

Pohlaví	Akaikeho informační kritérium										
	Model	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Muži	Coale-Kisker	-16,7	-3,0	-20,7	-10,2	-24,1	-11,2	-11,4	-32,1	1,3	2,8
	Gompertz	-18,4	-2,8	-25,4	-22,8	-26,1	-12,8	-11,7	-34,8	-3,7	-2,2
	Gompertz-Makeham	-16,9	-2,8	-24,9	-12,0	-24,3	-11,4	-10,2	-32,8	0,2	1,1
	Heligman-Pollard	-18,4	-7,2	-26,4	-30,3	-26,1	-14,0	-14,2	-31,6	-3,4	-0,3
	Kannisto	-15,7	-12,8	-26,2	-37,1	-23,7	-13,6	-16,2	-25,5	-2,6	2,5
	Thatcher	-14,5	-9,4	-23,3	-25,0	-23,1	-11,8	-13,3	-26,0	0,9	4,4
Pohlaví	Akaikeho informační kritérium										
	Model	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Muži	Coale-Kisker	-11,6	-12,0	-13,1	-5,8	-16,4	-19,5	-17,0	-9,4	-16,1	-19,2
	Gompertz	-11,1	-13,8	-7,4	-9,0	-16,1	-18,6	-16,8	-12,6	-17,1	-27,8
	Gompertz-Makeham	-9,9	-10,6	-10,4	-6,6	-14,1	-19,3	-16,6	-10,7	-15,0	-20,9
	Heligman-Pollard	-14,2	-18,7	-9,2	-14,2	-21,0	-23,2	-14,8	-16,0	-19,1	-32,6
	Kannisto	-17,7	-24,2	-11,8	-20,3	-27,3	-28,2	-12,7	-18,3	-19,9	-35,8
	Thatcher	-14,4	-18,2	-12,0	-14,8	-21,6	-25,4	-13,8	-15,9	-17,8	-29,0
Pohlaví	Akaikeho informační kritérium										
	Model	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Muži	Coale-Kisker	-1,6	-22,3	-12,1	-23,4	-5,5	-23,3	-28,9	-18,3	-40,2	-18,6
	Gompertz	-11,7	-28,8	-37,6	-39,9	-31,6	-58,3	-35,6	-39,8	-62,2	-30,6
	Gompertz-Makeham	-3,8	-24,5	-19,8	-30,2	-11,8	-32,8	-31,7	-24,2	-55,3	-23,2
	Heligman-Pollard	-16,7	-32,6	-43,1	-42,4	-39,2	-55,0	-32,8	-41,9	-53,1	-31,3
	Kannisto	-21,2	-35,7	-44,9	-41,1	-44,9	-48,0	-29,3	-42,4	-46,4	-30,1
	Thatcher	-13,9	-31,6	-37,2	-40,0	-25,5	-48,9	-31,8	-34,7	-66,0	-29,7

**Zdroj:** vlastní zpracování z výstupu DeRaSu (Burcin a kol., 2012) z dat poskytnutých ČSÚ (2017a).

**Pozn.:** červeně jsou označeny nejlepší hodnoty Akaikeho informačního kritéria v daném roce.

Na obrázku 27 jsou zobrazeny výsledky Akaikeho informačního kritéria pro jednotlivé vyrovnávací funkce u mužů. Z časové řady o 30 pozorování vychází v 16 případech nejlépe funkce Kannisto. Na základě tohoto kritéria jsou pro muže zvoleny úmrtnostní tabulky vyrovnané právě prostřednictvím funkce Kannisto.

**Obr. 28: Akaikeho informační kritérium při volbě vyrovnávací funkce u žen mezi lety 1986–2015**

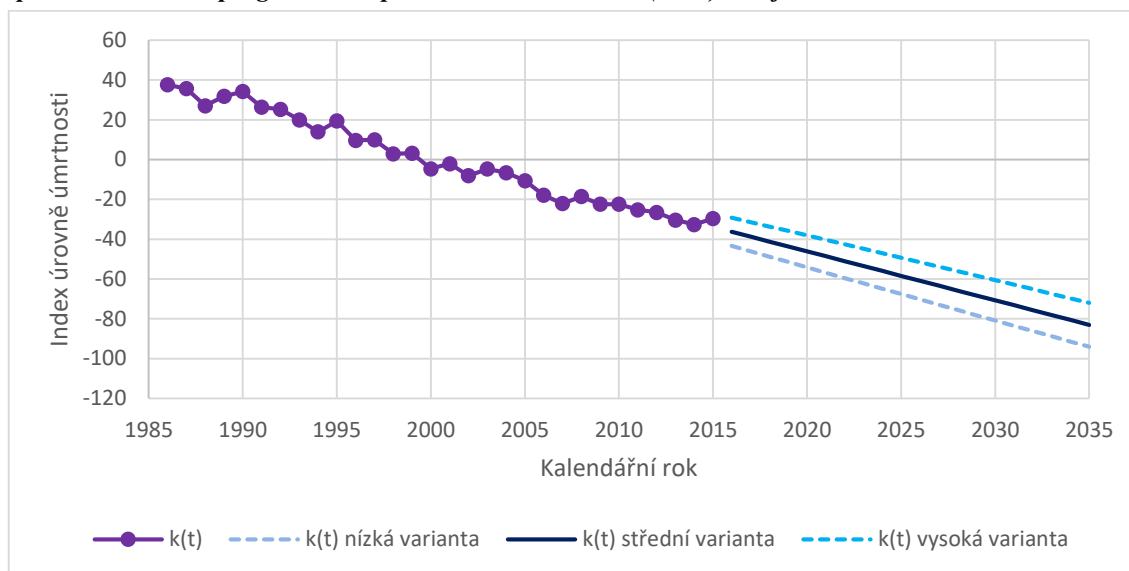
Pohlaví	Akaikeho informační kritérium										
	Model	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
Ženy	Coale-Kisker	-20,2	-1,3	-10,8	-5,6	-15,7	-24,8	-60,6	-4,7	-16,8	-21,5
	Gompertz	-18,2	-2,5	-4,5	-9,0	-12,1	-25,9	-31,3	-1,2	-11,3	-24,4
	Gompertz-Makeham	-18,1	-0,1	-6,1	-6,8	-12,5	-24,0	-44,9	-2,6	-12,4	-20,9
	Heligman-Pollard	-24,0	-8,8	-10,8	-13,6	-19,7	-39,4	-41,8	-7,0	-18,1	-34,4
	Kannisto	-27,9	-16,3	-18,9	-16,9	-29,3	-56,2	-57,4	-14,2	-26,3	-46,6
	Thatcher	-25,8	-11,4	-16,7	-14,2	-25,4	-48,0	-59,8	-12,3	-23,6	-39,2
Pohlaví	Akaikeho informační kritérium										
	Model	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Ženy	Coale-Kisker	-36,0	1,9	-17,0	-27,3	0,2	-11,9	-47,0	-23,5	-33,8	-27,3
	Gompertz	-24,3	-6,6	-6,7	-20,0	-1,4	-14,8	-51,1	-26,1	-30,8	-39,1
	Gompertz-Makeham	-32,1	-0,6	-9,0	-23,6	0,1	-12,0	-54,6	-24,1	-30,2	-30,0
	Heligman-Pollard	-30,9	-15,1	-13,5	-28,4	-8,6	-24,2	-54,8	-38,5	-41,8	-58,4
	Kannisto	-36,8	-25,5	-23,3	-39,5	-17,7	-35,7	-46,9	-52,9	-53,6	-71,0
	Thatcher	-36,8	-16,7	-21,5	-38,5	-13,2	-29,7	-42,9	-47,1	-49,5	-62,2
Pohlaví	Akaikeho informační kritérium										
	Model	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Ženy	Coale-Kisker	-53,5	-44,1	-30,2	-22,5	-11,5	-10,7	-10,9	-14,2	-35,4	-11,5
	Gompertz	-47,3	-33,3	-32,0	-24,6	-28,7	-29,7	-35,5	-43,4	-45,1	-37,3
	Gompertz-Makeham	-44,0	-35,4	-27,4	-20,9	-18,4	-16,1	-19,1	-22,6	-40,3	-20,9
	Heligman-Pollard	-65,9	-46,4	-39,9	-34,7	-39,1	-40,9	-48,7	-56,2	-52,9	-48,1
	Kannisto	-99,1	-62,1	-45,6	-45,3	-48,3	-53,1	-59,9	-64,7	-54,9	-54,0
	Thatcher	-82,1	-58,1	-41,5	-40,2	-38,0	-37,4	-44,3	-47,5	-52,1	-43,1

**Zdroj:** vlastní zpracování z výstupu DeRaSu (Burcin a kol., 2012) z dat poskytnutých ČSÚ (2017a).

**Pozn.:** červeně jsou označeny nejlepší hodnoty Akaikeho informačního kritéria v daném roce.

Podobným způsobem jako u mužů, je i u žen nutné zjistit, jaká funkce nejlépe aproximuje původní data. Z obrázku 28 vyplývá, že z časové řady o 30 pozorování vychází z 28 případů nejlépe rovněž funkce Kannisto, kde hodnotícím měřítkem je Akaikeho informační kritérium. Pro ženy jsou zvoleny úmrtnostní tabulky vyrovnané funkcí Kannisto a jejich tabulkové míry úmrtnosti tvoří stejným způsobem jako u mužů základ Lee-Carterova modelu úmrtnosti.

**Obr. 29:** Vývoj indexu úrovně úmrtnosti mužů mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno prostřednictvím  $ARIMA(0,1,1)$  s driftem

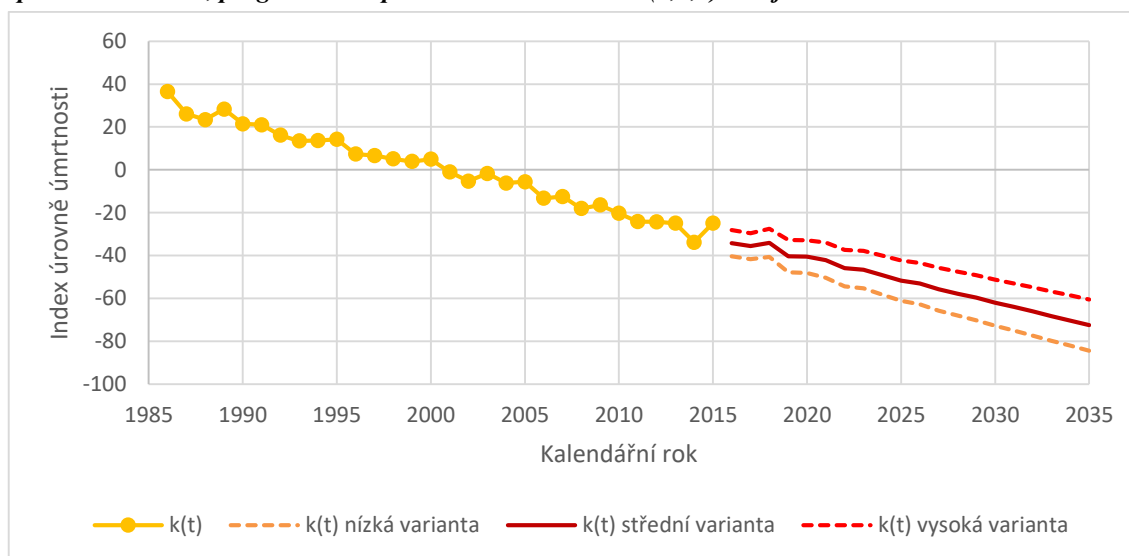


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $k(t)$  představuje index úrovně úmrtnosti z Lee-Carterova modelu úmrtnosti.

Prostřednictvím Lee-Carterova modelu úmrtnosti mužů se z vyrovnaných tabulkových měr úmrtnosti mužů odhadnou parametry tohoto modelu. Na obrázku 29 je zobrazena prognóza indexu úrovně úmrtnosti. Fialovou barvou jsou označeny odhadnuté hodnoty indexu úrovně úmrtnosti ze skutečných dat a od roku 2016 je vytvořena prognóza tohoto indexu s 95% spolehlivostí. Pro prognózu byl zvolen smíšený model  $ARIMA$  z důvodu, že dokáže identifikovat a odstranit trend, a tím se stává výsledný model stacionárním. Na základě příkazu (7.3) byl zvolen jako optimální smíšený model  $ARIMA(0,1,1)$  s driftem. Z toho vyplývá, že do modelu není zařazeno žádné zpoždění časové řady procesu AR. Již z obrázku 29 je patrný lineární trend, tento trend se odstraní první diferencí této časové řady. Do modelu je dále zařazena jedna zpožděná hodnota náhodného šoku procesu  $MA(1)$ . Drift představuje v modelu nenulovou střední hodnotu jednou diferencované časové řady. Tuto střední hodnotu je nutné z časové řady odstranit prostřednictvím centrování. Zásadou této skutečnosti se časová řada stává stacionární.

**Obr. 30:** Vývoj indexu úrovně úmrtnosti žen mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno prostřednictvím ARIMA(2,1,0) s driftem

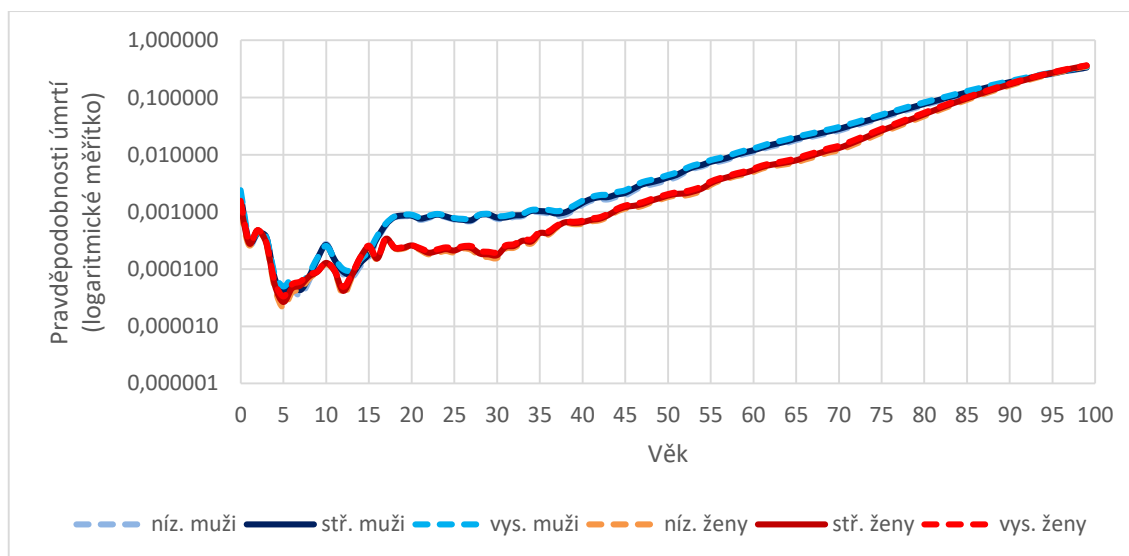


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $k(t)$  představuje index úrovně úmrtnosti z Lee-Carterova modelu úmrtnosti.

Stejným způsobem jako u mužů se i u žen odhadnou parametry Lee-Carterova modelu úmrtnosti z vyrovnaných tabulkových měr úmrtnosti, v tomto případě žen. Na obrázku 30 je zobrazena prognóza indexu úrovně úmrtnosti u žen se spolehlivostí 95 %. Oranžovou barvou jsou znázorněny odhadnuté hodnoty indexu úrovně úmrtnosti spočítané ze skutečných vyrovnaných tabulkových měr a od roku 2016 je zpracována prognóza tohoto indexu prostřednictvím modelu ARIMA(2,1,0) s driftem. Do časové řady indexu úrovně úmrtnosti jsou zařazeny dvě její zpoždění procesu AR(2), které se projevují na obrázku 30 zvlněním ve všech variantách. První diferencí se odstraní lineární trend a prostřednictvím centrování se časová řada stacionarizuje.

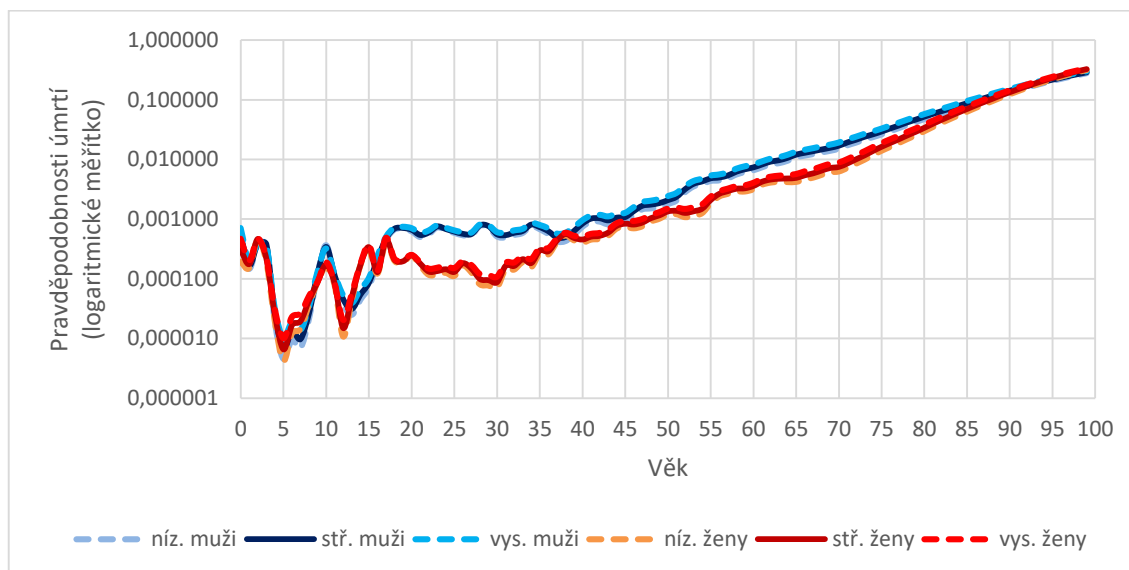
**Obr. 31:** Očekávaný věkový profil pravděpodobností úmrtí mužů a žen v roce 2016 se spolehlivostí 95 %



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Vzhledem ke skutečnosti, že data pro Jihočeský kraj jsou v této práci k dispozici do roku 2015, byla populační prognóza s 95% pravděpodobností zpracována od roku 2016. Z obrázku 31 je patrné, že v roce 2016 se profil pravděpodobností úmrtí v rámci variant téměř neměnil. Porovná-li se úmrtnost mužů a žen, tak do 15. roku se úmrtnost téměř nelišila, od 15. roku začínala výrazná nadúmrtnost mužů a v nejstarších věcích docházelo k vyrovnání těchto rozdílů.

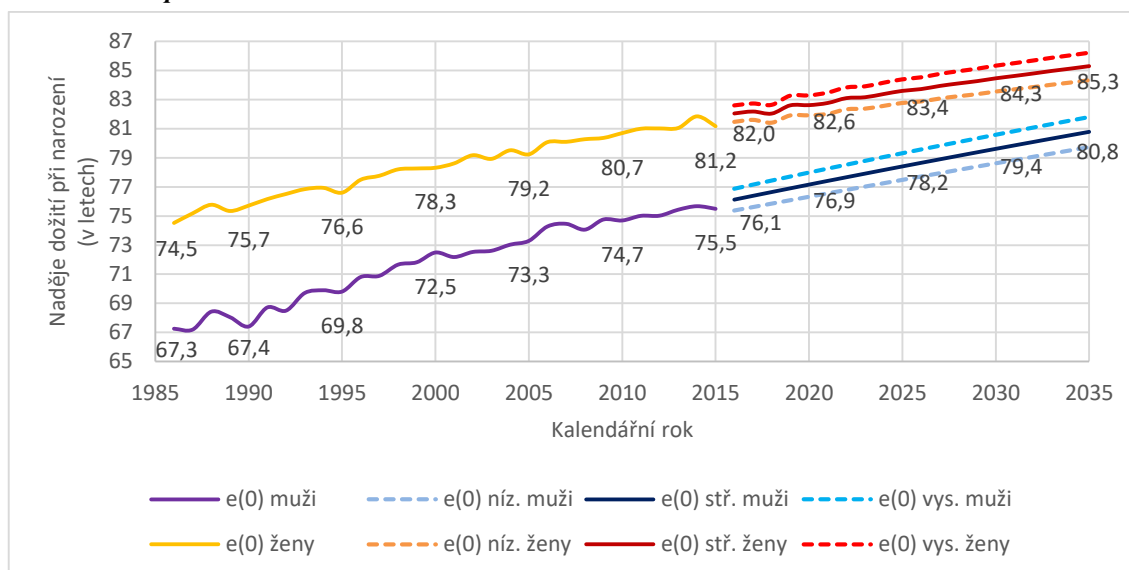
**Obr. 32: Očekávaný věkový profil pravděpodobností úmrtí mužů a žen v roce 2035 se spolehlivostí 95 %**



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Populační prognóza Jihočeského kraje byla zpracována do roku 2035. Na obrázku 32 jsou zobrazeny očekávané pravděpodobnosti úmrtí mužů a žen podle věku v roce 2035 ve třech variantách. Již při prognóze indexu úrovně úmrtnosti mužů z obrázku 29 a žen z obrázku 30 je zřejmé, že změny v rámci variant nebudou příliš velké. Tato skutečnost se potvrdila i na obrázku 32. Porovná-li se úmrtnost z roku 2016 (Obr. 31) a úmrtnost v roce 2035 (Obr. 32), tak vzhledem k zachování stejného měřítka obou grafů, je na první pohled zřetelný výrazný pokles pravděpodobností úmrtí. Tento pokles je výraznější u mužů nežli u žen. Dále lze porovnat rozdíly v profilu úmrtnosti mezi muži a ženami. Očekává se, že do 15. věku se výrazněji odstraní rozdíly mezi muži a ženami při porovnání prognózy z roku 2016 oproti roku 2035. Mezi věky 15 až 35 se předpokládá rozdíl mezi úmrtností mužů a žen nejvyšší. Lze si to vysvětlit tím, že ženy jsou všeobecně opatrnější a úmrtnost na vnější příčiny se tím bude u nich snižovat. Přibližně od věku 35 let se očekává, že se úmrtnost mužů a žen bude sbližovat. Tuto skutečnost lze vysvětlit tím, že ženy budou přejímat nezdravé návyky mužů a muži budou více dbát o své zdraví. Podle Čermákové (2009, s. 18) je zpomalení poklesu úmrtnosti žen oproti mužům způsobeno emancipací žen a vysvětluje to na příkladu kouření, kdy dříve bylo kouření vnímáno jako nátura mužů a po druhé světové válce při přebírání mužských pozic ženami dochází k značnému nárůstu kuřáček. Tento horší způsob života se neprojevuje ve věku, kdy člověk začne kouřit, ale právě ve vyšších věcích.

**Obr. 33:** Vývoj naděje dožití chlapců a dívek při narození mezi lety 1986–2015 a jejich prognóza do roku 2035 se spolehlivostí 95 %



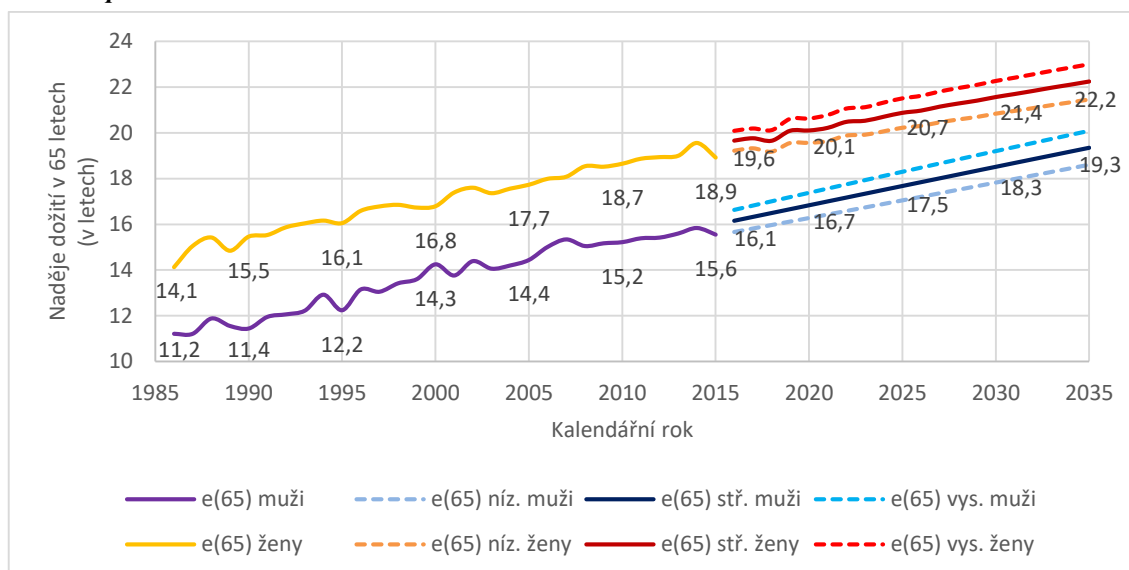
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** e(0) představuje naději dožití při narození.

Na obrázku 33 je zpracována prognóza naděje dožití chlapců a dívek při narození v Jihočeském kraji se spolehlivostí 95 %. Oranžovou barvou je zobrazen skutečný vývoj naděje dožití při narození dívek a fialovou barvou je znázorněn skutečný vývoj naděje dožití při narození chlapců. Již z tohoto vývoje je patrná konvergence naděje dožití chlapců a dívek při narození a lze očekávat, že se tato skutečnost promítne i do její prognózy. Tmavě modrou barvou je zobrazena střední varianta prognózy naděje dožití chlapců, tmavě červená barva symbolizuje střední variantu prognózy naděje dožití dívek, čerchovanou barvou jsou vyobrazeny meze spolehlivosti pro naději dožití s 95% spolehlivostí. Od roku 1986 do roku 2015 vzrostla naděje dožití při narození u dívek o 6,7 let. Do roku 2035 se podle střední varianty předpokládá její nárůst ještě o 4 roky. U chlapců je nárůst naděje dožití při narození o 8,2 roků mezi roky 1986 až 2015. Do roku 2035 se podle střední varianty očekává její nárůst ještě o 5,3 roku. Rozdíl naděje dožití při narození mezi muži a ženami je v roce 1986 přibližně 7,2 roku. V roce 2015 je tento rozdíl 5,7 roků a podle střední varianty se v roce 2035 očekává tento rozdíl 4,5 roku, kdy vyšší naděje dožití při narození zůstává stále u žen.

Stejným způsobem jako lze analyzovat vývoj naděje dožití při narození, lze studovat také naději dožití v jiných věcích. Pro ilustraci je na obrázku 34 uveden vývoj naděje dožití mužů a žen ve věku 65 let, včetně prognózy do roku 2035. Podle střední varianty prognózy se očekává, že od roku 2015 do roku 2035 naroste naděje dožití mužů ve věku 65 o 3,7 let, u žen se předpokládá toto zvýšení nižší, a to o 3,3 roku. Lze očekávat, že se rozdílná naděje dožití mužů a žen ve věku 65 let bude dále sbližovat. V roce 2015 je tento rozdíl 3,3 roku a v roce 2035 se předpokládá snížení tohoto rozdílu na 2,9 roku.

**Obr. 34: Vývoj naděje dožití mužů a žen v 65 letech mezi lety 1986–2015 a jejich prognóza do roku 2035 se spolehlivostí 95 %**



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $e(65)$  představuje naději dožití v 65 letech.

V tabulce 10 jsou zobrazeny předpokládané příspěvky k rozdílu hodnot naděje dožití při narození mezi roky 2015 a 2035 zvláště pro muže a zvláště pro ženy ve věkových skupinách 0–14, 15–64 a 65 a více let. Rok 2015 je posledním rokem, ke kterému jsou k dispozici data pro tuto diplomovou práci, pro rok 2035 je zvolena střední varianta prognózy naděje dožití při narození z obrázku 33. Metodický postup této dekompozice popisuje Arriaga (1984, s. 86–89). Základem této dekompozice je rozdíl dvou nadějí dožití při narození a jejich rozklad na přímý efekt, nepřímý efekt a interakci při zahrnutí různé intervalové délky života. Výpočet se provádí z úmrtnostních tabulek ve zmíněných dvou letech a z jejich tabulkových funkcí  $l_x$ ,  $T_x$  a  $e_x$ , kde  $l_x$  představuje tabulkový počet dožívajících se přesného věku  $x$ ,  $T_x$  značí celkový počet let zbývajících na dožití tabulkové populace v dokončeném věku  $x$  a  $e_x$  je střední délka života v přesném věku  $x$ .

**Tab. 10: Očekávané příspěvky ke změně naděje dožití při narození mezi lety 2015–2035 ve věkových skupinách 0–14, 15–64 a 65 a více let**

Pohlaví	Věková skupina	Přímý efekt	Nepřímý efekt	Interakce	Celkem	Celkem (v %)
Ženy	0–14	0,02	0,10	0,01	0,13	2,94
	15–64	0,30	0,57	0,11	0,98	22,26
	65+	3,28	–	–	3,28	74,80
	<b>Celkem</b>	3,60	0,67	0,12	4,38	–
	<b>Celkem (v %)</b>	82,17	15,19	2,64	–	–
Muži	0–14	0,06	0,25	0,02	0,33	6,20
	15–64	0,67	0,94	0,23	1,84	34,77
	65+	3,12	–	–	3,12	59,03
	<b>Celkem</b>	3,85	1,19	0,25	5,29	–
	<b>Celkem (v %)</b>	72,79	22,50	4,72	–	–

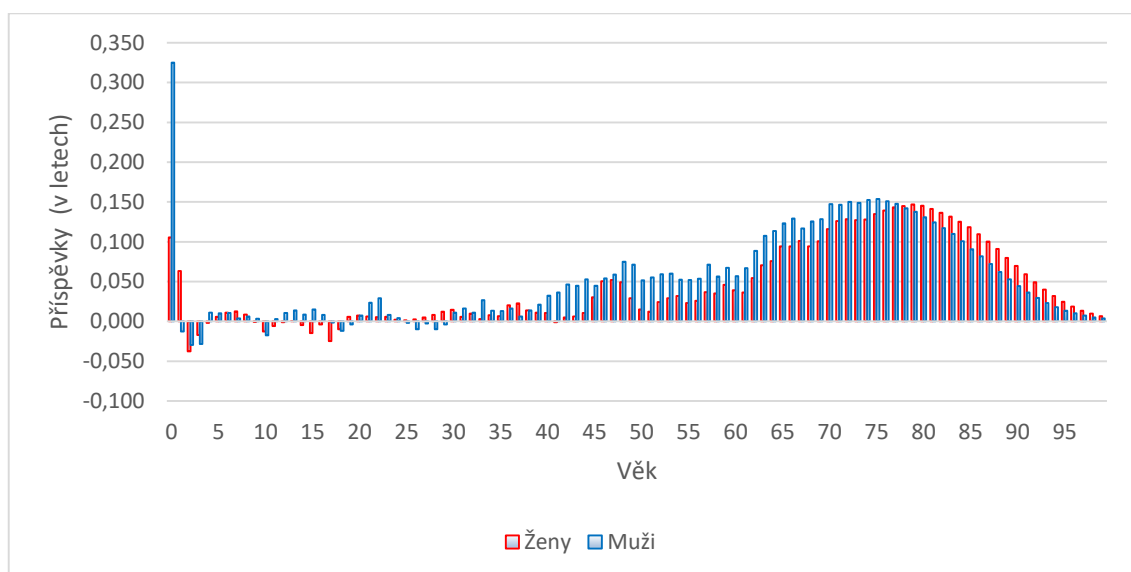
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Z tabulky 10 je patrné, že nejvyšší příspěvky k rozdílu hodnot celkové naděje dožití při narození mezi roky 2015 a 2035 se očekávají ve věkové skupině 65 a více let, kde se předpokládá



nárůst těchto příspěvků u žen až o 3,28 let a u mužů o 3,12 let. Příspěvky v ostatních věkových skupinách již budou s největší pravděpodobností nižší. Celkově se očekává rychlejší nárůst příspěvků k rozdílu hodnot naděje dožití při narození mezi roky 2015 a 2035 u mužů nežli u žen. Ve věkové skupině 0–14 je tempo růstu těchto příspěvků téměř trojnásobně vyšší u mužů než u žen a ve věkové skupině 15–64 přibližně dvojnásobně rychlejší ve prospěch mužů. U obou pohlaví vychází, že rozdíl v naději dožití při narození mezi roky 2015 až 2035 je převážně způsobený změnou intenzit popsaných v tabulce přímým efektem než věkovou strukturou charakterizovanou nepřímým efektem.

**Obr. 35:** Očekávané příspěvky ke změně naděje dožití při narození podle jednotek věku mezi lety 2015–2035



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** pro rok 2035 je uvažována střední variantu prognózy.

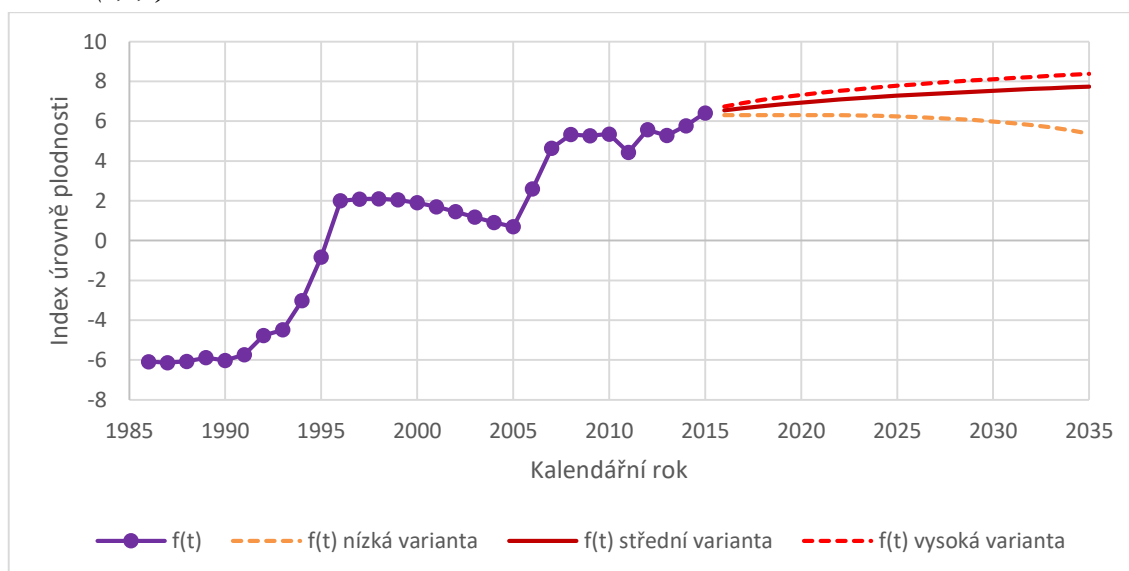
Na obrázku 35 je prezentován graf předpokládaných příspěvků k rozdílu hodnot naděje dožití při narození mezi lety 2015–2035 podle jednotek věku pro obě pohlaví. Stejně jako v předchozím případě rok 2035 označuje střední variantu prognózy a rok 2015 označuje skutečná data. Pro výpočet byla zvolena jednorozměrná dekompozice podle Pressata, kterou metodicky popsal Vallin a Caselli (2006, s. 211). Princip této dekompozice spočívá v tom, že součet příspěvků ke změně naděje dožití při narození, v tomto případě mezi dvěma roky, se musí rovnat rozdílu celkové naděje dožití mezi těmito roky. Samotný výpočet se provádí z úmrtnostních tabulek ve zmíněných dvou letech a z jejich tabulkových funkcí  $l_x$  a  $e_x$ , kde  $l_x$  představuje tabulkový počet dožívajících se přesného věku  $x$  a  $e_x$  je střední délka života v přesném věku  $x$ . Z grafu vyplývá, že nejvyšší příspěvky k růstu naděje dožití při narození lze předpokládat u mužů výrazným poklesem kojenecké úmrtnosti. Tato skutečnost je zapříčiněna tím, že v roce 2015 byla kojenecká úmrtnost u mužů výrazně vyšší, tzn. 4,7 ‰ než v předchozích letech, kdy například v roce 2014 byla pouze 2,4 ‰. To se promítá u mužů ve zdánlivě vyšších příspěvcích k růstu naděje dožití při narození. U žen se očekávají tyto přírůstky nejvyšší okolo věku 79 let. Obecně platí u obou pohlaví vyšší příspěvky k růstu naděje dožití při narození v kojeneckém období a v nejvyšších

věcích. Při porovnání těchto příspěvků mezi muži a ženami, lze očekávat, že přibližně od věku 40 let dojde mezi roky 2015–2035 k jejich rychlejšímu nárůstu u mužů. Tento rychlejší růst je s největší pravděpodobností způsoben určitou rezervou, která se vytvořila v minulosti kvůli rychlejšímu poklesu úmrtnosti u žen oproti mužům. Muži mají zásluhou snižování úmrtnosti v posledních letech větší potenciál k rychlejšímu růstu naděje dožití. U žen se oproti mužům očekávají výraznější příspěvky k růstu naděje dožití přibližně od věku 80 let. Záporné hodnoty v příspěvcích k růstu naděje dožití při narození, které se projevují v dětském období, jsou způsobeny nulovými hodnotami kvocientů úmrtnosti v roce 2015 v tomto dětském období.

## 7.2 Dílčí prognóza plodnosti zpracovaná s využitím Lee-Carterova modelu plodnosti

Před zpracováním dílčí prognózy plodnosti je zapotřebí nejprve vypočítat míry plodnosti podle věku z dlouhé výchozí časové řady počtu narozených. Tyto míry tvoří základ Lee-Carterova modelu plodnosti. Následně je nutné odhadnout parametry tohoto modelu a vytvořit prognózu indexu úrovně plodnosti. Tento index je dále modelován prostřednictvím smíšeného modelu ARIMA. Popsána zde bude dílčí prognóza plodnosti Jihočeského kraje s 95% spolehlivostí. Dílčí prognóza plodnosti s 75% spolehlivostí je zpracována v příloze 7 a 8.

**Obr. 36:** Vývoj indexu úrovně plodnosti mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno exponenciální a zpětnou logaritmickou transformací prostřednictvím  $ARIMA(0,2,1)$



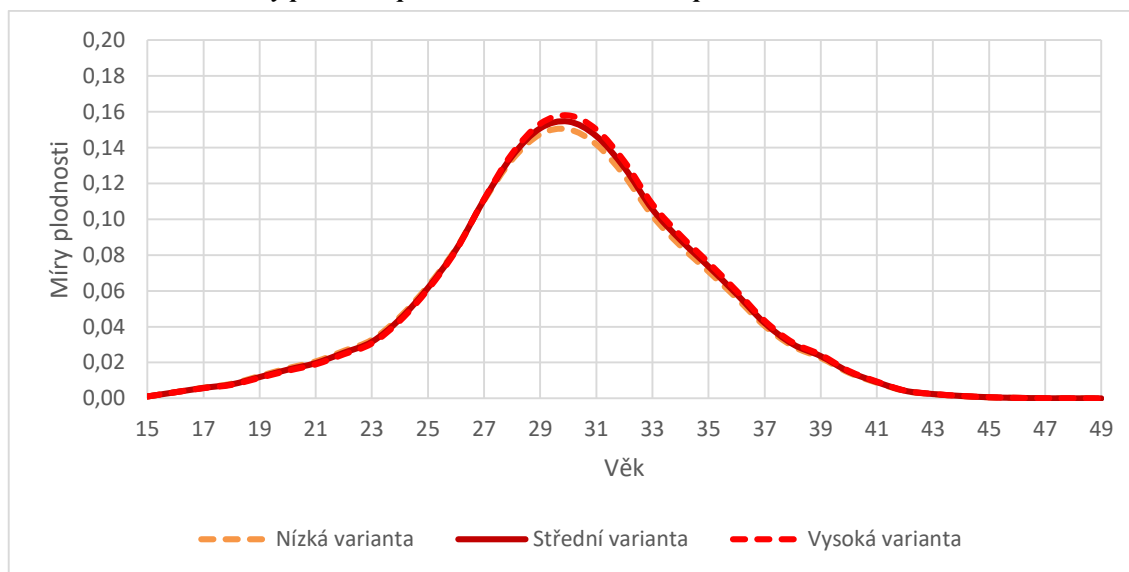
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $f(t)$  představuje index úrovně plodnosti z Lee-Carterova modelu plodnosti.

Na obrázku 36 je znázorněna prognóza indexu úrovně plodnosti s 95% spolehlivostí. Fialovou barvou jsou vyznačeny odhadnuté hodnoty tohoto indexu ze skutečných dat. Vzhledem k nereálným mezím spolehlivosti při použití klasického modelu ARIMA bylo nutné časovou řadu indexu úrovně plodnosti transformovat. Pro tuto transformaci byla zvolena exponenciální funkce.

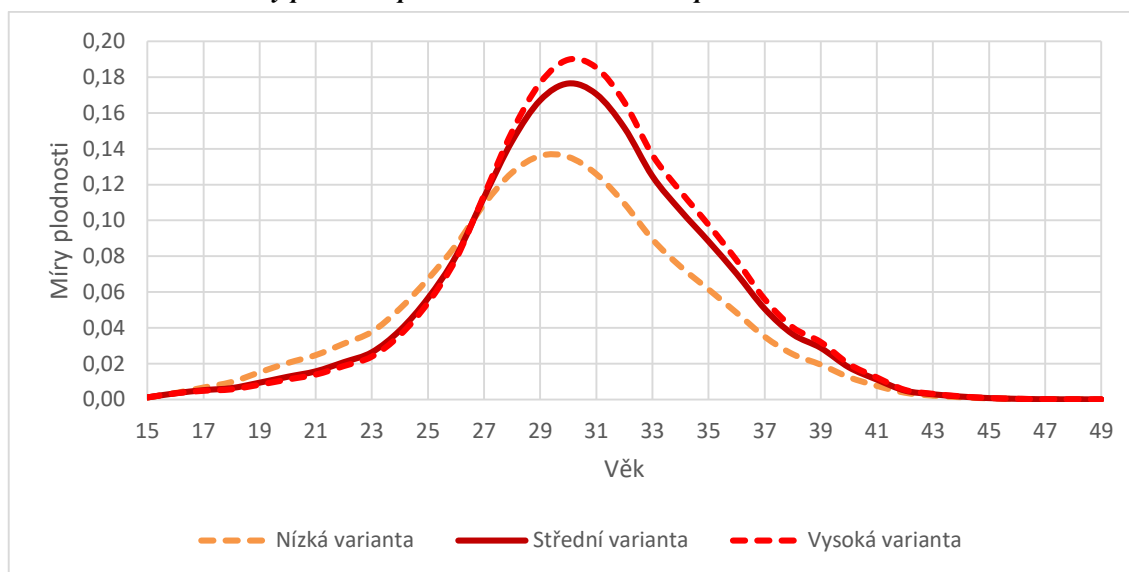
Na takto upravenou časovou řadu byl aplikován příkaz (7.3) a jako nejvhodnější model byl zvolen smíšený ARIMA(0,2,1) model, prostřednictvím kterého se tato časová řada stacionarizuje. Do modelu není zařazena žádná zpožděná hodnota takto upravené časové řady procesu AR. Časová řada je dvakrát diferencovaná a použije se jedna hodnota náhodného šoku procesu MA(1). Tímto způsobem vzniklé odhady se zpětně transformují logaritmováním a výsledná prognóza indexu úrodnosti je od roku 2016 vyznačena na obrázku 36.

**Obr. 37:** Očekávané míry plodnosti podle věku v roce 2016 se spolehlivostí 95 %



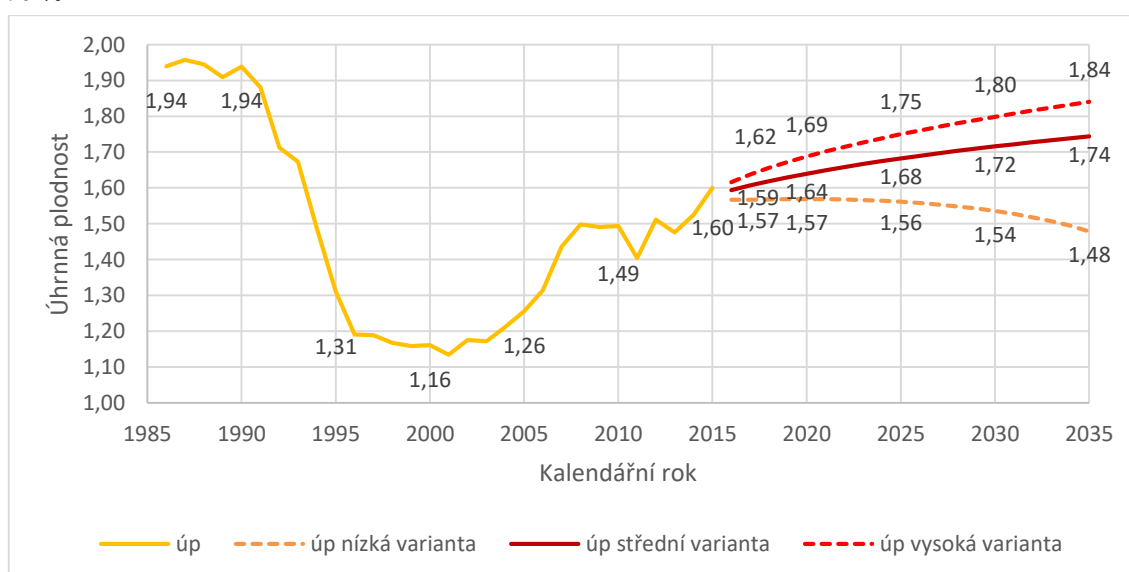
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Vzhledem k tomu, že údaje za počty narozených jsou dostupné do roku 2015, bylo nutné prognózu plodnosti zpracovat od roku 2016. Na obrázku 37 jsou patrné míry plodnosti podle věku s 95% spolehlivostí v Jihočeském kraji, které se očekávaly v roce 2016. Protože se jedná o počátek prognózy, předpokládaly se rozdíly mezi jednotlivými variantami téměř nepatrné. Podle střední varianty prognózy začala plodnost od 15 let pozvolna narůstat a od 23 let narůstala strměji. Modus plodnosti byl přibližně ve věku 30 let a od tohoto věku byl strmý pokles plodnosti až do věku 43 let. Ve vyšším věku byla plodnost prakticky nulová. Průměrný věk matky při narození dítěte byl 30,6 roků.

**Obr. 38:** Očekávané míry plodnosti podle věku v roce 2035 se spolehlivostí 95 %

**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Obrázek 38 znázorňuje prognózu měr plodnosti podle věku v roce 2035 se spolehlivostí 95 % pro Jihočeský kraj. Vysoká varianta i střední varianta předpokládají zvýšení měr plodnosti podle věku oproti roku 2016 (Obr. 37). Pouze nízká varianta očekává nižší úroveň plodnosti než v roce 2016. Průměrný věk matky při narození dítěte se očekává v roce 2035 podle nízké varianty 30,1 let, podle střední varianty 31,0 let a podle vysoké varianty 31,3 let. V příloze 16 jsou zobrazeny očekávané změny věkového profilu měr plodnosti v rámci střední varianty prognózy mezi roky 2016 až 2035. Předpokládá se postupný nárůst intenzit plodnosti v tomto období a již zmíněný mírný posun průměrného věku matek z 30,6 let v roce 2016 na 31,0 let v roce 2035.

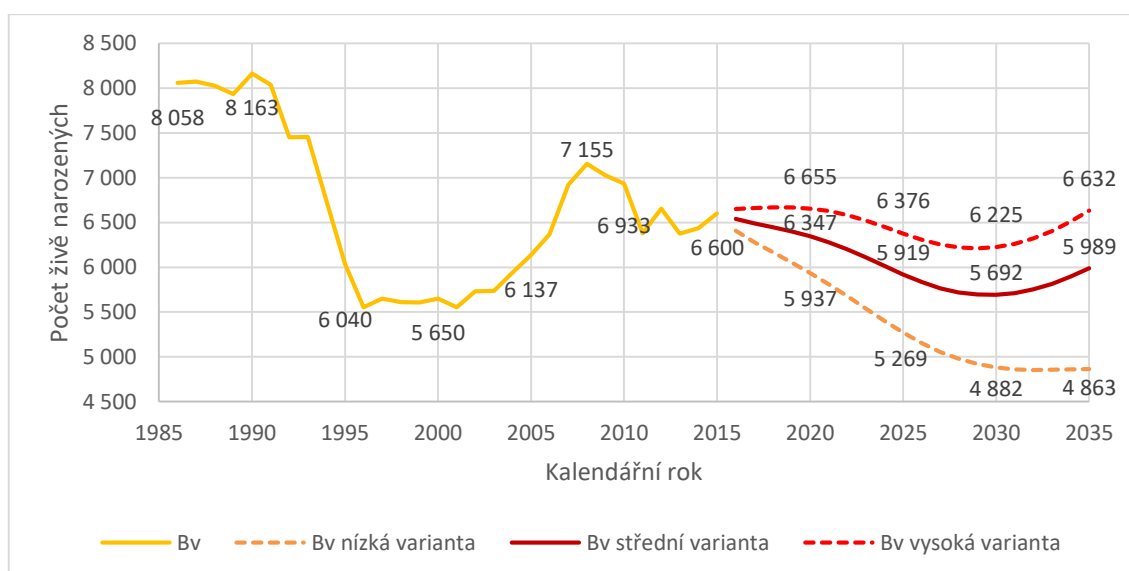
**Obr. 39:** Vývoj úhrnné plodnosti mezi lety 1986–2015 a její prognóza do roku 2035 se spolehlivostí 95 %

**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** úp představuje úhrnnou plodnost.

Obrázek 39 zobrazuje vývoj úhrnné plodnosti v Jihočeském kraji včetně její prognózy do roku 2035 se spolehlivostí 95 %. Oranžovou barvou jsou znázorněny skutečné hodnoty úhrnné plodnosti. Vývoj úhrnné plodnosti zaznamenal během minulých třiceti let výrazné změny, kdy se její úroveň z hodnot 1,94 v roce 1986 propadla v roce 2000 na hodnotu 1,16. V posledních letech dochází k podpoře porodnosti, tato podpora má pozitivní účinek na zvýšení počtu narozených dětí. Výše úhrnné plodnosti je ale stále nízká, protože nedosahuje hranice prosté reprodukce, tzn. v průměru 2,10 dítěte na jednu ženu. Pokud je úhrnná plodnost pod touto hranicí, lze očekávat úbytek počtu obyvatel přirozenou měrou. Od roku 2016 je zpracována prognóza úhrnné plodnosti, kdy střední varianta je znázorněna tmavě červenou barvou a meze spolehlivosti jsou označeny čerchovanými čarami. Vysoká i střední varianta předpokládají nárůst úhrnné plodnosti. Nízká varianta předpokládá její pokles.

**Obr. 40:** Vývoj počtu narozených mezi lety 1986–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** Bv představuje počty živě narozených.

Obrázek 40 ukazuje vývoj počtu živě narozených v Jihočeském kraji během let 1986 až 2015 a od roku 2016 je zobrazena jejich prognóza s 95 % spolehlivostí. Za posledních 30 let byly nejvyšší počty narozených v 80. letech 20. století, naopak v 90. letech nastává výrazný pokles počtu narozených způsobený odkladem mateřství do vyššího věku. Podrobněji se odkladem mateřství do vyššího věku zabývá Zdražilová (2016). Tento odklad způsobuje nárůst počtu narozených od roku 2000 s vrcholem v roce 2008 a od té doby dochází k pozvolnému poklesu počtu narozených. Tento pokles předpokládají všechny varianty prognóz mezi roky 2016 až 2030. Nejhorší situace nastává při nízké variantě prognózy, kdy počty narozených, při úrovni úhrnné plodnosti 1,48 v roce 2035, by poklesly pod 5000 dětí. Naopak vysoká varianta i střední varianta, předpokládají pokles do roku 2030 a mírný nárůst počtu narozených do roku 2035. Tento nárůst je způsobený především věkovou strukturou, kdy se do věku maximální plodnosti dostávají silné ročníky narozené po roce 2000.

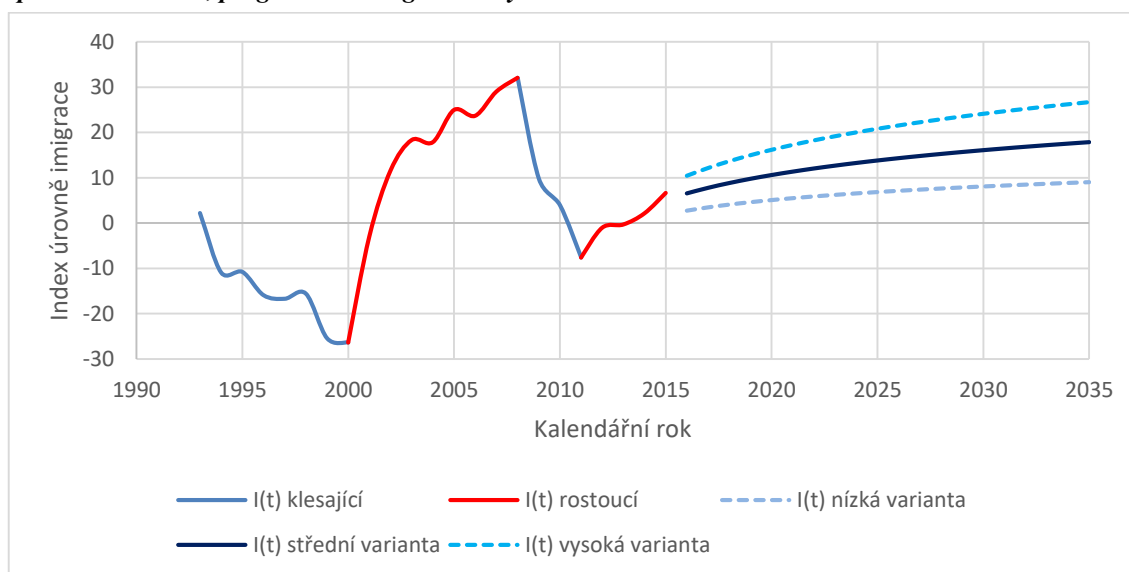
## 7.3 Dílčí prognóza migrace

Tato podkapitola se věnuje zpracování a prezentaci výsledků dílčí prognózy migrace v Jihočeském kraji z dlouhé výchozí časové řady. Prognóza migrace se skládá z dílčí prognózy imigrace a dílčí prognózy emigrace. Prostředkem k vytvoření této prognózy je Lee-Carterův model. Na závěr této kapitoly budou popsány výsledky migračního salda.

### 7.3.1. Dílčí prognóza imigrace s využitím Lee-Carterova modelu imigrace

Dílčí prognóza imigrace vychází z dlouhé výchozí časové řady počtu přistěhovalých mužů a žen v Jihočeském kraji, která je modelována Lee-Carterovým modelem imigrace. Vzhledem ke skutečnosti, že počty přistěhovalých nezávisí na věkové struktuře, je imigrace analyzována skutečnými počty přistěhovalých. Po odhadu parametrů Lee-Carterova modelu imigrace je zapotřebí dále modelovat index úrovně imigrace. Protože délka výchozí časové řady je pouze 23 pozorování, nelze užít smíšený model ARIMA. K modelování tohoto indexu jsou zvoleny trendové funkce, které lze použít i pro kratší časové řady. Diskutována zde bude prognóza imigrace Jihočeského kraje se spolehlivostí 95 %. Prognóza imigrace s 75% spolehlivostí je zobrazena v příloze 9 a 10.

**Obr. 41:** Vývoj indexu úrovně imigrace mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem



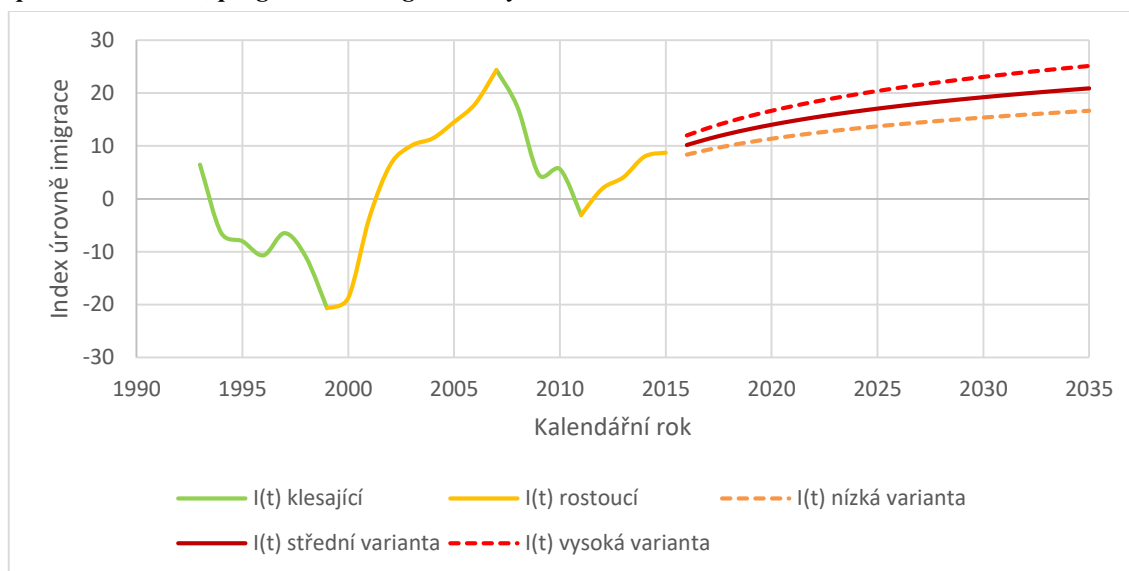
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $I(t)$  představuje index úrovně imigrace z Lee-Carterova modelu imigrace.

Na obrázku 41 je ukázán vývoj indexu úrovně imigrace mužů během období 1993 až 2015 a od roku 2016 je sestavena prognóza tohoto indexu s 95% spolehlivostí. Tuto prognózu nebylo možné vzhledem k délce výchozí časové řady modelovat prostřednictvím smíšeného modelu ARIMA. Při důkladné analýze odhadnutého indexu úrovně imigrace bylo zjištěno, že se tato časová řada skládá ze čtyř částí, a to ze dvou klesajících částí znázorněných modře a ze dvou rostoucích částí zobrazených červeně. Tento vývoj indexu úrovně imigrace lze přisuzovat

ekonomickým podmínkám České republiky. Od roku 2000 dochází k ekonomickému růstu České republiky a tím i počet přistěhovalých do Jihočeského kraje narůstá. V roce 2008 dochází ke světové krizi, která se projevuje i v České republice, kde dochází k výrazným rozpočtovým škrtnutím. Tento vývoj má vliv i na počet přistěhovalých, kdy jejich počet výrazně klesá. Od roku 2010 dochází k oživování ekonomiky a od roku 2011 dochází i k růstu počtu přistěhovalých v Jihočeském kraji. Vzhledem k tomuto vývoji je index úrovně imigrace modelován od roku 2011, kdy se předpokládá růst počtu přistěhovalých jako v letech před krizí. Křivka indexu úrovně imigrace od roku 2000 až 2008 má přibližně logaritmický trend. Pokud by nenastala krize, pak by tento trend s největší pravděpodobností pokračoval i po roce 2008. Stejný vývoj indexu úrovně imigrace lze tedy očekávat i od roku 2011. Logaritmický trend je odhadnut z hodnot tohoto indexu mezi lety 2011 až 2015. Od roku 2016 je prostřednictvím odhadnutého logaritmického trendu vytvořena prognóza indexu úrovně imigrace.

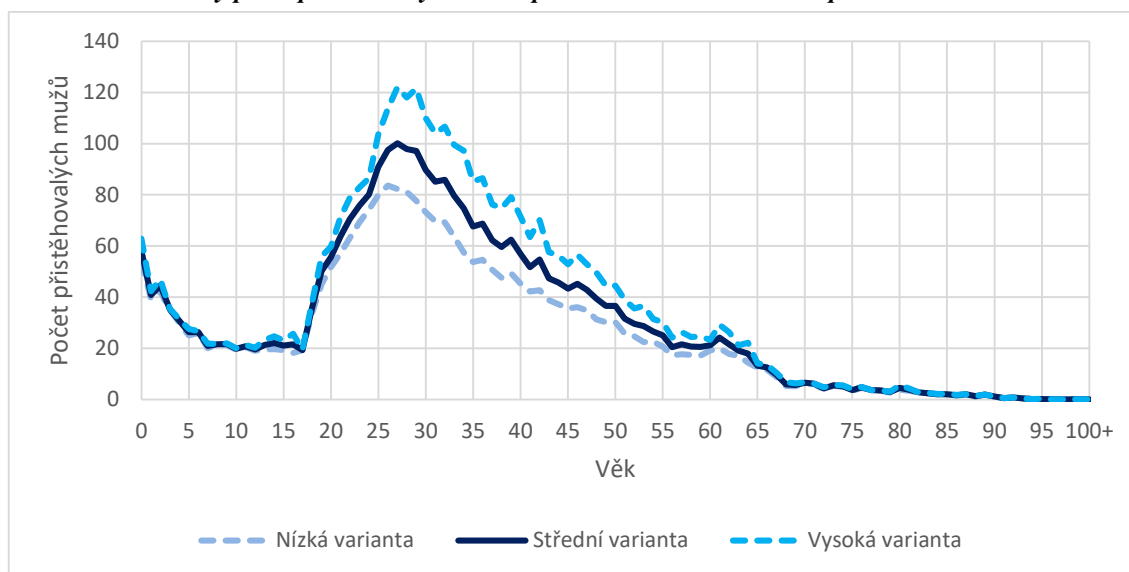
**Obr. 42:** Vývoj indexu úrovně imigrace žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

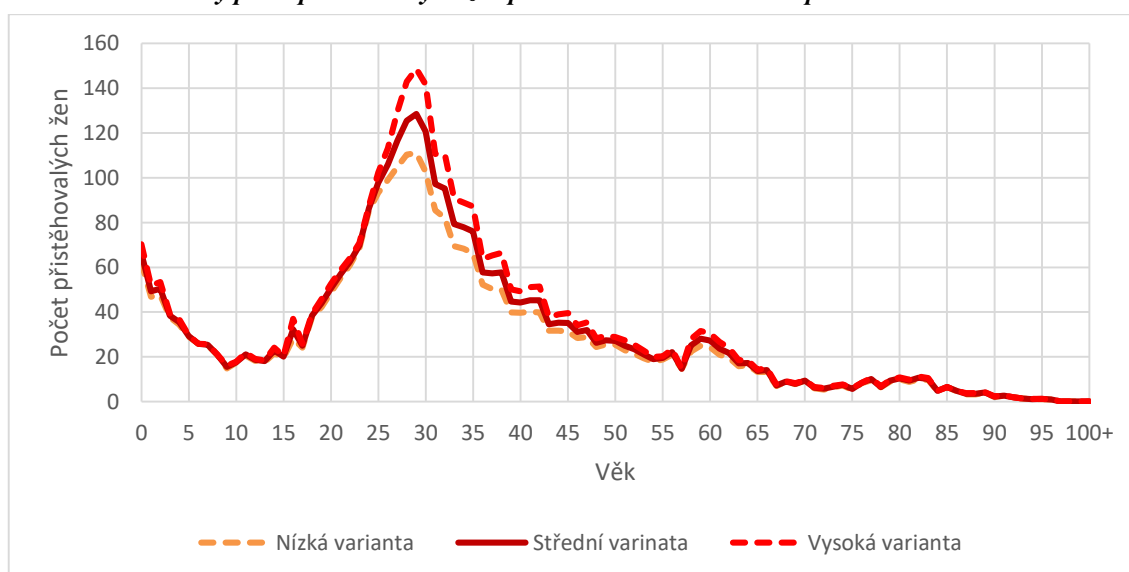
**Pozn.:**  $I(t)$  představuje index úrovně imigrace z Lee-Carterova modelu imigrace.

Index úrovně imigrace mužů má přibližně stejný průběh i u žen viz obrázek 42. Z tohoto důvodu byla aplikována stejná metoda odhadu tohoto indexu i u žen, tzn. prostřednictvím indexu úrovně imigrace mezi roky 2011 a 2015 se odhadnou parametry logaritmického trendu. Od roku 2016 dochází k vytvoření prognózy tohoto indexu prostřednictvím logaritmického trendu. Porovná-li se prognóza indexu úrovně imigrace mužů (Obr. 41) a žen (Obr. 42), tak u mužů je šířka intervalu spolehlivosti větší. Je to způsobeno tím, že časová řada tohoto indexu kopíruje méně přesně logaritmický trend.

**Obr. 43:** Očekávaný počet přistěhovaných mužů podle věku v roce 2035 se spolehlivostí 95 %

**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

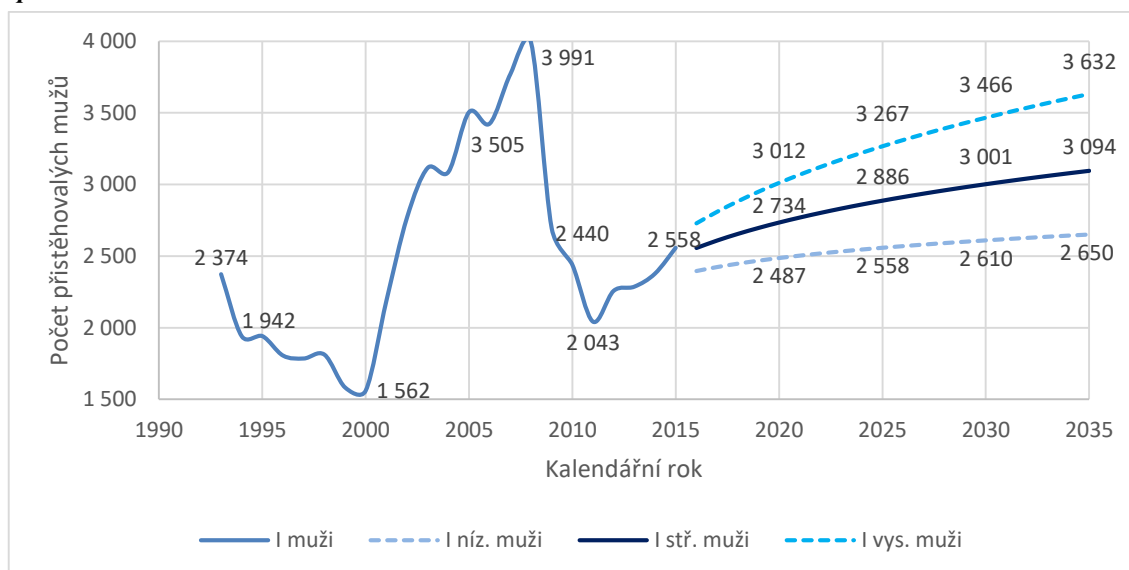
Obrázek 43 zachycuje očekávanou strukturu počtu přistěhovaných mužů podle jednotek věku v roce 2035. Stejným způsobem je znázorněna předpokládaná struktura počtu přistěhovaných žen na obrázku 44, kdy plnou čarou je zobrazena střední varianta prognózy a čerchované čáry symbolizují 95% interval spolehlivosti kolem této střední varianty. Očekává se přibližně stejná struktura přistěhovaných mužů a žen, kdy nejvíce přistěhovaných se očekává ve věkové skupině 25 až 30 let. Z toho lze usoudit, že se jedná především o migraci za prací. Druhé výrazné zvýšení počtu přistěhovaných nastává při narození, kdy lidé, kteří se stěhují za prací ve věkové kategorii 25 až 30 let, také nejčastěji v těchto věcích zakládají rodinu. Poslední mírný lokální vzestup se očekává v 60 letech, kdy toto stěhování může být spojené s přechodem do postproduktivního věku.

**Obr. 44:** Očekávaný počet přistěhovaných žen podle věku v roce 2035 se spolehlivostí 95 %

**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).



**Obr. 45:** Vývoj počtu přistěhovaných mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %

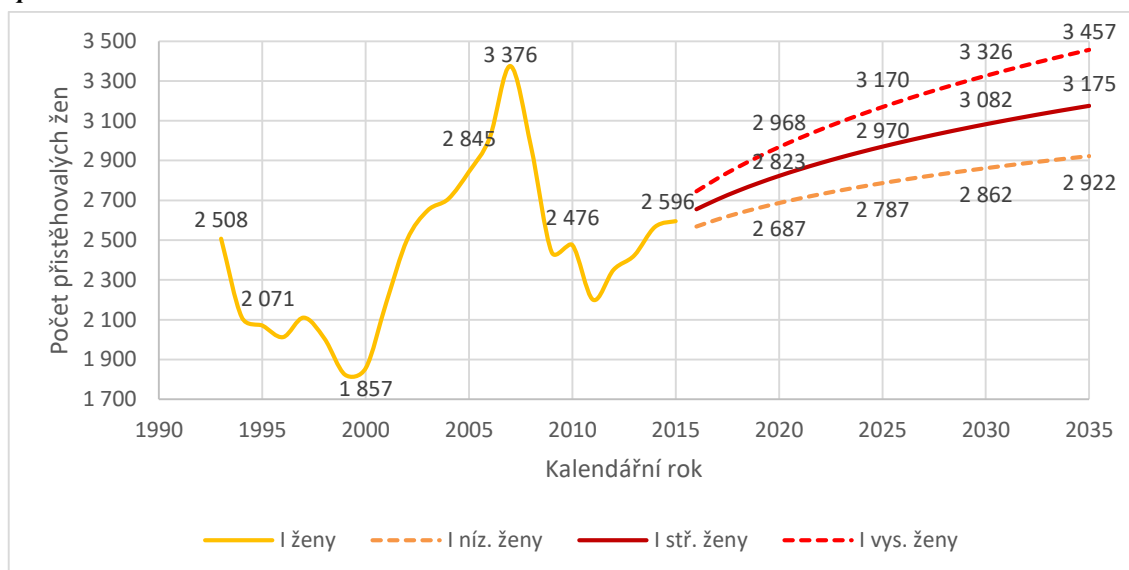


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** I představuje počet přistěhovaných.

Na obrázku 45 je zpracován vývoj celkového počtu přistěhovaných mužů, na obrázku 46 je zobrazen vývoj celkového počtu přistěhovaných žen a od roku 2016 je vytvořena jejich prognóza. Trend vývoje počtu přistěhovaných za období 1993 až 2015 byl již popsán v úvodu této podkapitoly, při popisování techniky konstrukce prognózy indexu úrovně imigrace. Na tomto místě budou analyzovány výsledky prognózy počtu přistěhovaných mužů a žen. Podle střední varianty se předpokládá, že celkově bude více přistěhovaných žen než mužů. Je to způsobeno tím, že z výchozích dat ve většině let převládaly přistěhované ženy nad přistěhovanými muži, tento trend lze očekávat i do budoucna. Může to být zapříčiněno tím, že ženy se více stěhují za svými protějšky než muži. Vzhledem k vyšší variabilitě odhadu počtu přistěhovaných mužů to ale nelze říci jednoznačně. Předpokládá se, že počet přistěhovaných mužů bude narůstat, ale tempo tohoto růstu se bude postupně snižovat. Toto tvrzení dokládá střední varianta počtu přistěhovaných mužů, která očekává růst mezi lety 2015 až 2020 o 176, mezi lety 2020 až 2025 o 152, mezi lety 2025 až 2030 o 115 a mezi lety 2030 až 2035 o 93 přistěhovaných mužů. Stejný vývoj lze předpokládat i u žen, kde je pětiletý nárůst mezi roky 2015 až 2020 o 227, mezi roky 2020 až 2025 o 147, mezi roky 2025 až 2030 o 112 a mezi roky 2030 až 2035 o 93 přistěhovaných žen. Z toho vyplývá, že se tempo nárůstu počtu přistěhovaných žen bude s největší pravděpodobností také snižovat.

**Obr. 46:** Vývoj počtu přistěhovaných žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

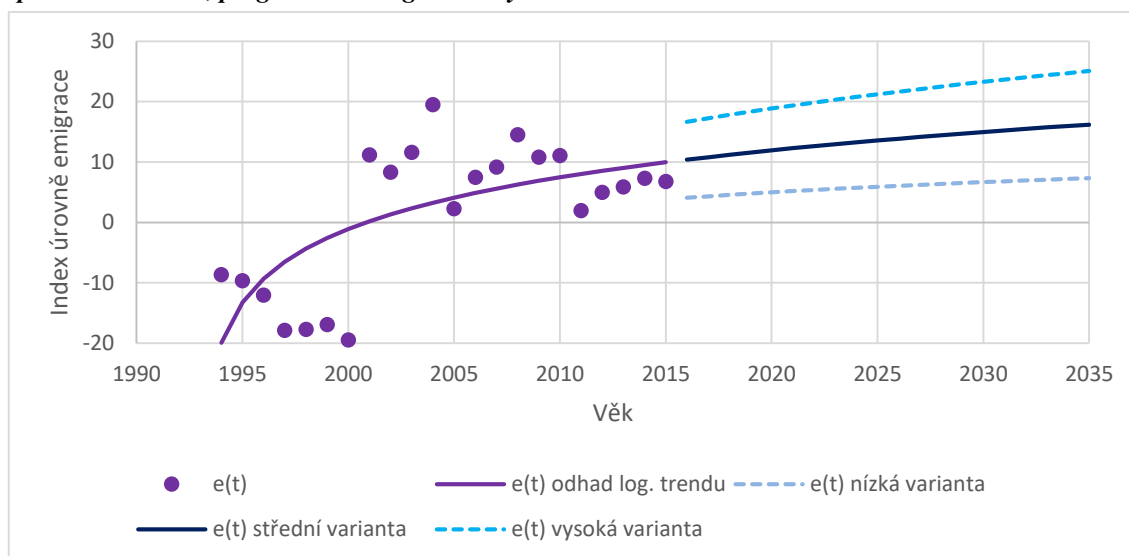
**Pozn.:** I představuje počty přistěhovaných.

### 7.3.2. Dílčí prognóza emigrace s využitím Lee-Carterova modelu emigrace

Při zpracování dílčí prognózy imigrace nezávisely počty přistěhovaných na věkové struktuře, u emigrace je to ale jiné. Čím více lidí žije v dané oblasti, tím je větší potenciál, že se z této oblasti někdo vystěhuje. Z tohoto vyplývá, že struktura a velikost populace mají velký vliv na počet vystěhovaných, z tohoto důvodu budou analyzovány míry vystěhování a nikoli skutečné počty vystěhovaných. Tyto míry poté vstupují do Lee-Carterova modelu emigrace. Vzhledem k tomu, že časovou řadu měř emigrace tvoří pouze 23 pozorování, nelze aplikovat smíšený model ARIMA a stejně jako u dílčí prognózy imigrace, se i zde vývoj indexu úrovně emigrace bude analyzovat prostřednictvím trendových funkcí. Na tomto místě bude popsána dílčí prognóza emigrace Jihočeského kraje s 95% spolehlivostí. V příloze 11 a 12 je uvedena tato prognóza s 75% spolehlivostí.

Obrázek 47 vyjadřuje vývoj indexu úrovně emigrace mužů včetně jeho prognózy se spolehlivostí 95 %. Výchozí data tohoto indexu vykazují přibližně logaritmický trend, který byl zvolen i k predikci. Problém ale nastává v roce 1993, kdy hodnota indexu emigrace mužů extrémně vybočuje z tohoto trendu. Z tohoto důvodu byla tato hodnota označena jako odlehlé pozorování a byla odstraněna z časové řady tohoto indexu. Na obrázku jsou fialovými tečkami označeny odhadnuté hodnoty indexu úrovně emigrace z Lee-Carterova modelu emigrace. Fialovou čarou je znázorněn logaritmický trend, který je statisticky významný na pětiprocentní hladině spolehlivosti. Statisticky jsou významné jak t-testy, tak i celkový F test. Od roku 2016 je vytvořena prognóza, kde střední varianta je zobrazena tmavě modrou plnou čarou a meze spolehlivosti jsou ukázány čerchovanou čarou s 95% spolehlivostí.

**Obr. 47:** Vývoj indexu úrovně emigrace mužů mezi lety 1994–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem

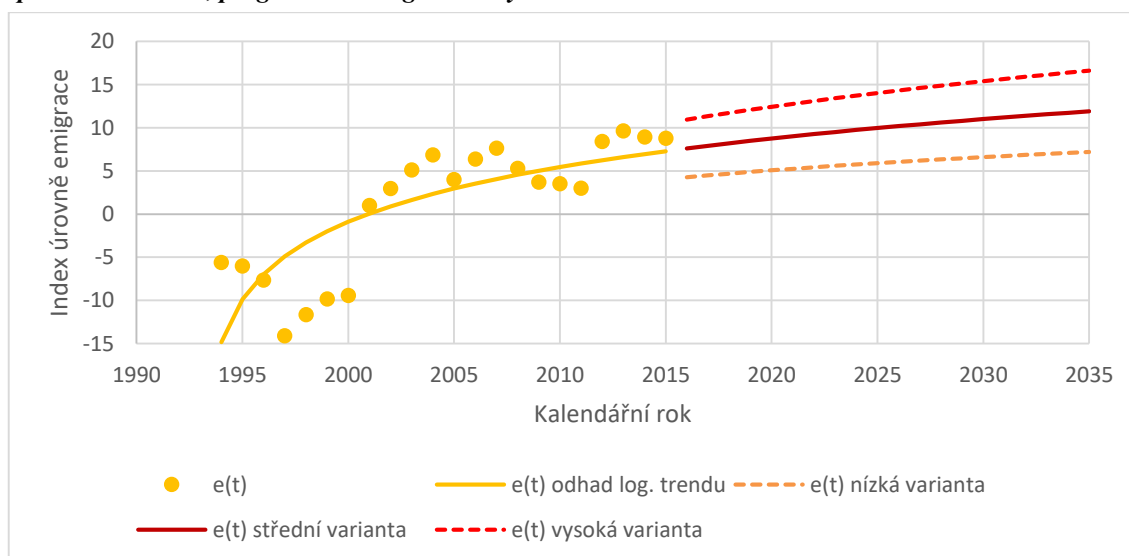


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $e(t)$  představuje index úrovně emigrace z Lee-Carterova modelu emigrace, log. značí logaritmický.

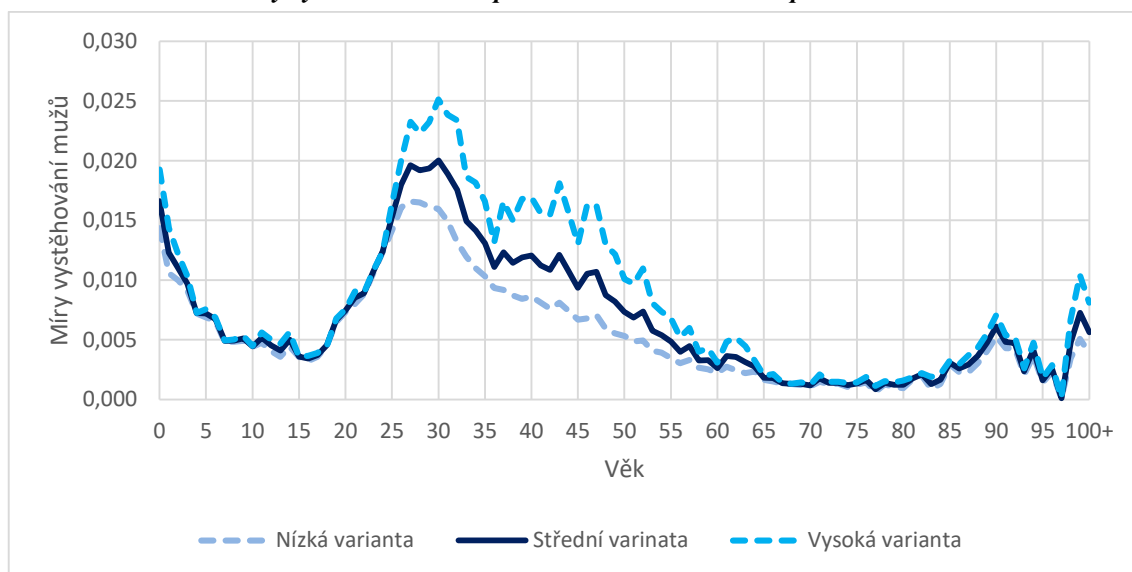
Vývoj indexu úrovně emigrace žen je prezentován na obrázku 48 včetně následné prognózy do roku 2035. Stejně jako u indexu úrovně emigrace mužů bylo zvoleno i u žen technické zpracování prognózy prostřednictvím logaritmického trendu. Vzhledem k vývoji tohoto indexu bylo odstraněno jeho odlehlé pozorování z roku 1993, která vybočovalo z tohoto trendu. Výsledné t-testy a celkový F test jsou významné. Vývoj indexu úrovně emigrace žen je velmi podobný s vývojem tohoto indexu u mužů. K rozdílnému patří odlišná variabilita obou odhadnutých souborů dat, která se poté projevuje i v prognóze širším intervalem spolehlivosti.

**Obr. 48:** Vývoj indexu úrovně emigrace žen mezi lety 1994–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %, prognózováno logaritmickým trendem



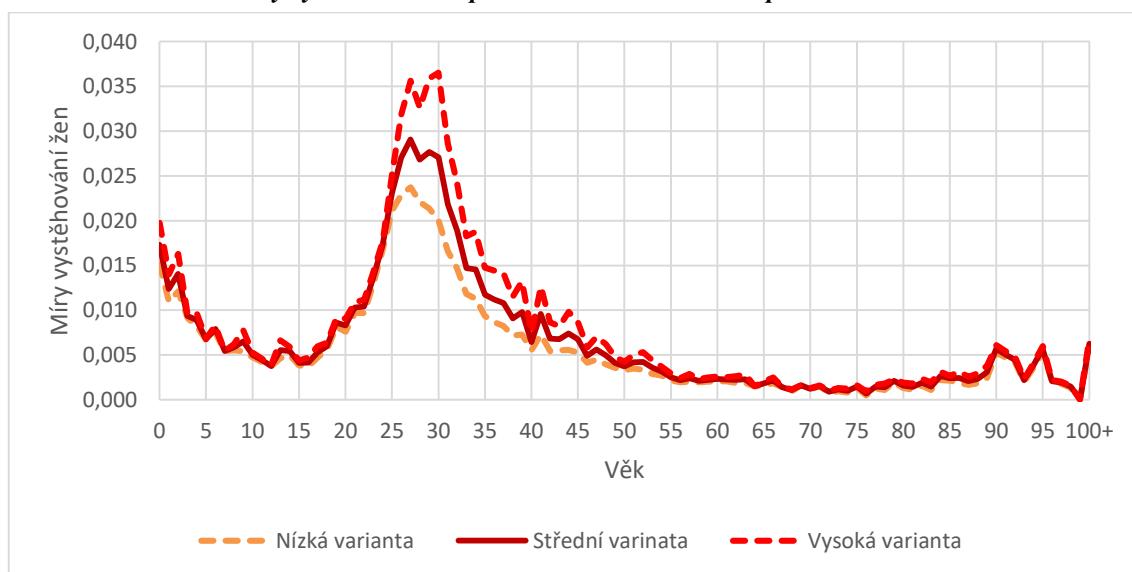
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:**  $e(t)$  představuje index úrovně emigrace z Lee-Carterova modelu emigrace, log. značí logaritmický.

**Obr. 49: Očekávané míry vystěhování mužů podle věku v roce 2035 se spolehlivostí 95 %**

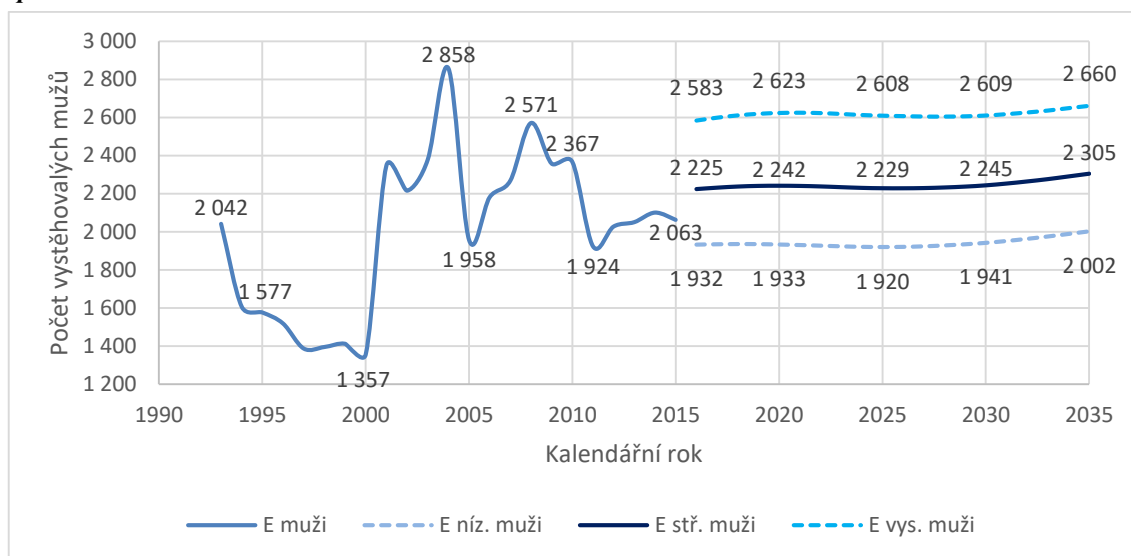
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Očekávaný vývoj měr vystěhování mužů v roce 2035 je zobrazen na obrázku 49 a vývoj měr vystěhování žen je v tomto roce zpracován na obrázku 50. Z obou grafů je patrné, že nejvyšších hodnot měr vystěhování je dosahováno při narození a ve věku 25 až 30 let. Může to být způsobeno tím, že ženy a muži se stěhují za prací a odchází z Jižních Čech včetně svých dětí. U žen se předpokládá pokles měr vystěhování po 30. roce strmější, naopak u mužů se očekává tento pokles pozvolnější. To může být způsobeno nižšími pracovními příležitostmi v Jihočeském kraji. Mírný nárůst v nejvyšších věcích je způsobený malým počtem osob v tomto věku. Již při analýze indexu úrovně emigrace mužů a žen byla zdůrazňována vyšší variabilita tohoto indexu u mužů, která se projevuje i v měrách vystěhování širším intervalem spolehlivosti kolem střední varianty.

**Obr. 50: Očekávané míry vystěhování žen podle věku v roce 2035 se spolehlivostí 95 %**

**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Obr. 51:** Vývoj počtu vystěhovaných mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %

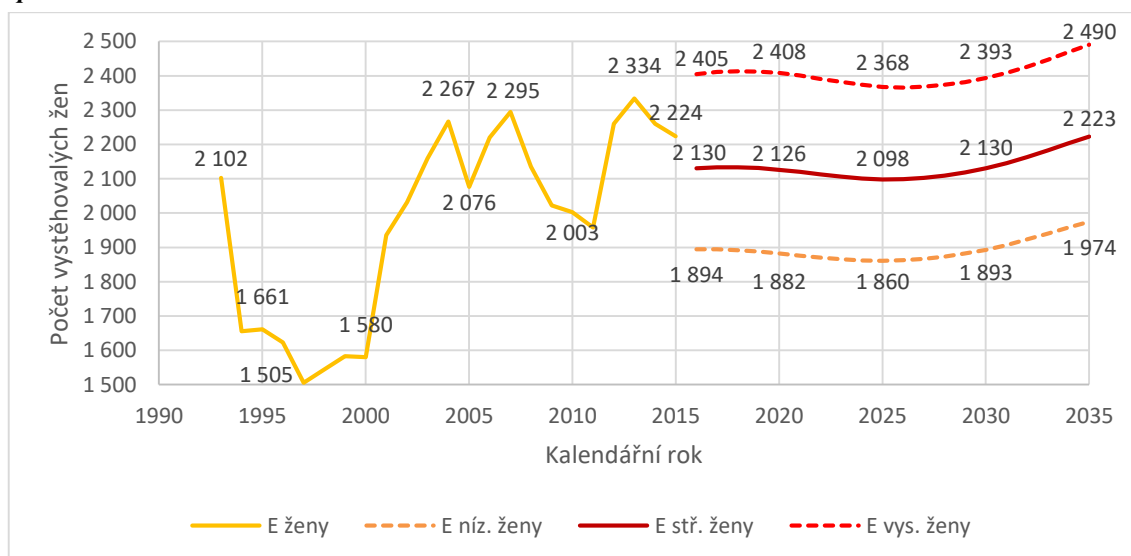


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** E představuje počet vystěhovaných.

Na obrázku 51 se nachází minulý vývoj počtu vystěhovaných mužů včetně jejich prognózy s 95 % spolehlivostí do roku 2035 a podobný vývoj je i u žen představený na obrázku 52. V roce 1993 byla úroveň vystěhování na vyšších hodnotách a následně mezi roky 1994 až 2000 nastal výrazný pokles jejich počtu. V roce 2001 došlo k výraznému růstu počtu vystěhovaných a jejich počet následně kolísá přibližně kolem hodnoty 2200. Podle prognózy se předpokládá, že počet vystěhovaných podle střední varianty poroste pouze mírně od roku 2016 do roku 2035 přibližně o 80 mužů a o 93 žen. Střední varianta rovněž předpokládá, že se vystěhuje více mužů než žen.

**Obr. 52:** Vývoj počtu vystěhovaných žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %



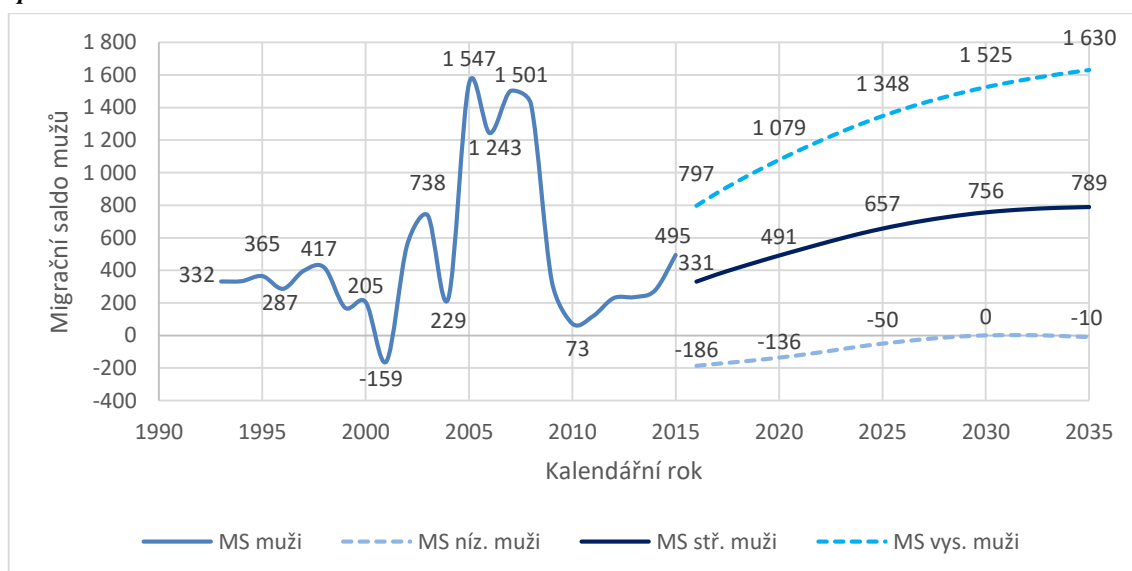
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** E představuje počet vystěhovaných.

### 7.3.3. Výsledné migrační saldo mužů a žen

Poslední podkapitola při analýze migrace je věnována popisu vývoje celkového migračního salda, které se vypočítá jako rozdíl mezi dílčí prognózou imigrace a dílčí prognózou emigrace. Na obrázku 53 je zobrazen vývoj migračního salda mužů v Jihočeském kraji a od roku 2016 je zpracována prognóza tohoto salda. Migrační saldo mezi lety 1993 až 2014 náhodně kolísá přibližně kolem hodnoty 300 mužů ročně. Mezi lety 2005 až 2008 dochází k velkému přílivu počtu přistěhovalých mužů, a to se projevuje i ve výrazném zvýšení migračního salda. Následný propad po roce 2008 je způsobený následky ekonomické krize. Od roku 2016 se očekává podle střední varianty růst migračního salda mužů z hodnot 331 v roce 2016 na hodnotu 789 v roce 2035. Vysoká varianta předpokládá razantnější růst až na hodnotu migračního salda 1630. Naopak nízká varianta očekává úbytek o 10 mužů způsobený migrací v roce 2035.

**Obr. 53:** Vývoj migračního salda mužů mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %

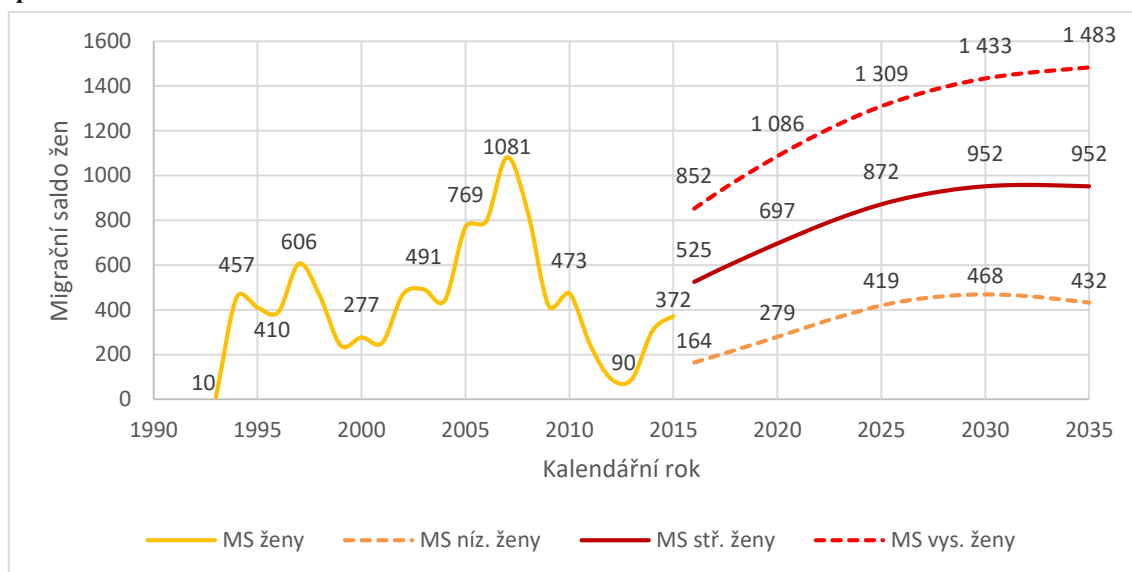


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** MS představuje migrační saldo.

Na obrázku 54 se nachází vývoj migračního salda žen Jihočeského kraje včetně jeho prognózy s 95% spolehlivostí. Nejnižší úroveň migračního salda žen byla v roce 1993, poté dochází k jeho nárůstu. Mezi lety 1993 až 2004 kolísá úroveň migračního salda žen přibližně kolem hodnoty 400. Od roku 2005 dochází k výraznému nárůstu tohoto salda a tento stav přetrvává až do roku 2008, kdy dochází kvůli krizi k výraznému poklesu migračního salda žen. Toto saldo postupně narůstá až v posledních letech. Do budoucna se očekává zachování tohoto růstu, kdy střední varianta uvažuje migrační saldo žen v roce 2035 na úrovni 952 žen, nízká varianta očekává toto saldo pouze na hodnotě 432 žen a vysoká varianta předpokládá navýšení až na 1483 žen.

**Obr. 54: Vývoj migračního salda žen mezi lety 1993–2015 a jeho prognóza do roku 2035 se spolehlivostí 95 %**



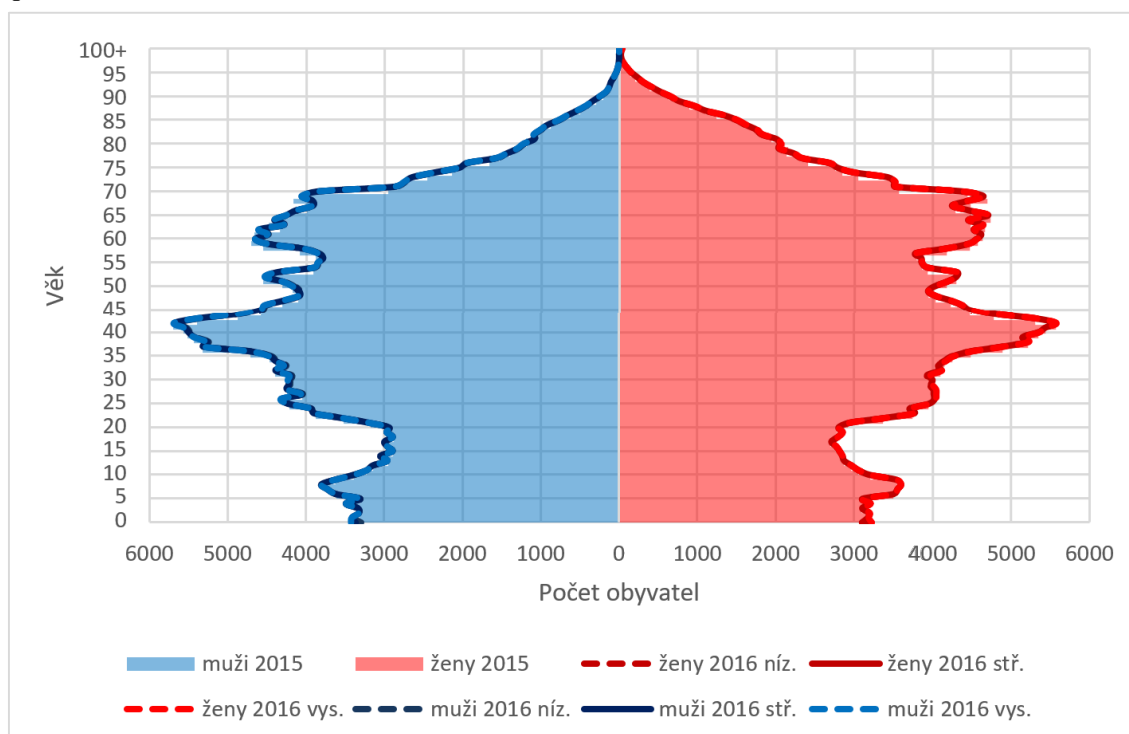
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** MS představuje migrační saldo.

#### 7.4 Populační prognóza počtu a věkové struktury mužů a žen v Jihočeském kraji zpracovaná z dílčích prognóz úmrtnosti, plodnosti a migrace prostřednictvím Lee-Carterových modelů

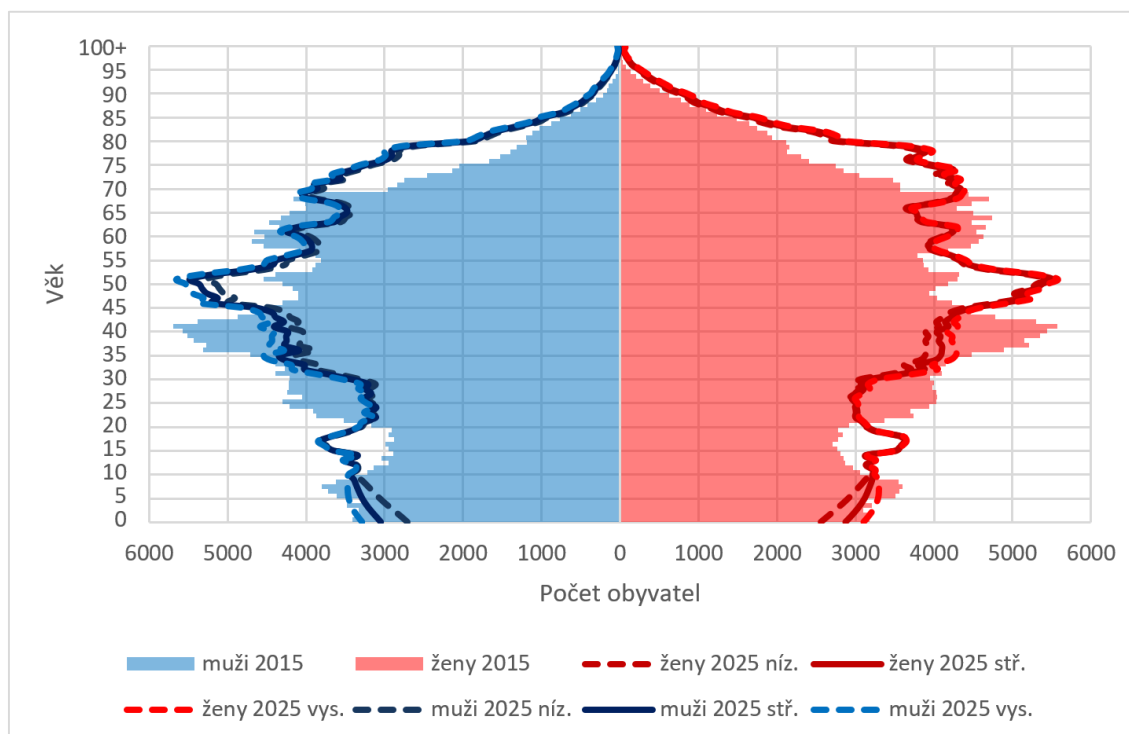
Tato podkapitola pojednává o výsledcích populační prognózy Jihočeského kraje, která byla zpracována z dílčích prognóz prostřednictvím analýzy dlouhé časové řady a byla vytvořena prostřednictvím kohortně komponentní metody s migrací s dvoufázovým zahrnutím imigrace a emigrace do modelu. Na obrázku 55 je zpracována prognóza struktury obyvatel Jihočeského kraje, která se očekávala v roce 2016. Jedná se o velmi krátký časový úsek populační prognózy a z tohoto důvodu byly rozdíly mezi variantami minimální. Výraznější posun se očekává k 31. 12. v roce 2025 oproti roku 2015 (Obr. 56), kdy struktura počtu mužů a žen s největší pravděpodobností zůstane přibližně stejná jako v roce 2015, pouze se posune v čase. Nejvyšší změny v rámci variant se očekávají v počtu narozených a ve věku 30 až 40 let u žen a 30 až 50 let u mužů. U počtu narozených dětí je to způsobené rozdílnou úrovní předpokládané úhrnné plodnosti v jednotlivých variantách. Podrobněji bylo téma zmíněno při zpracování dílčí prognózy plodnosti v podkapitole 7.2. Rozdíl u mužů a žen v produktivním věku je způsobený převážně různou úrovní imigrace, která byla rozebrána v podkapitole 7.3.1.

**Obr. 55: Porovnání prognózy věkové struktury mužů a žen k 31. 12. 2016 oproti 31. 12. 2015 s 95% spolehlivostí**



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Obr. 56: Porovnání prognózy věkové struktury mužů a žen k 31. 12. 2025 oproti 31. 12. 2015 s 95% spolehlivostí**

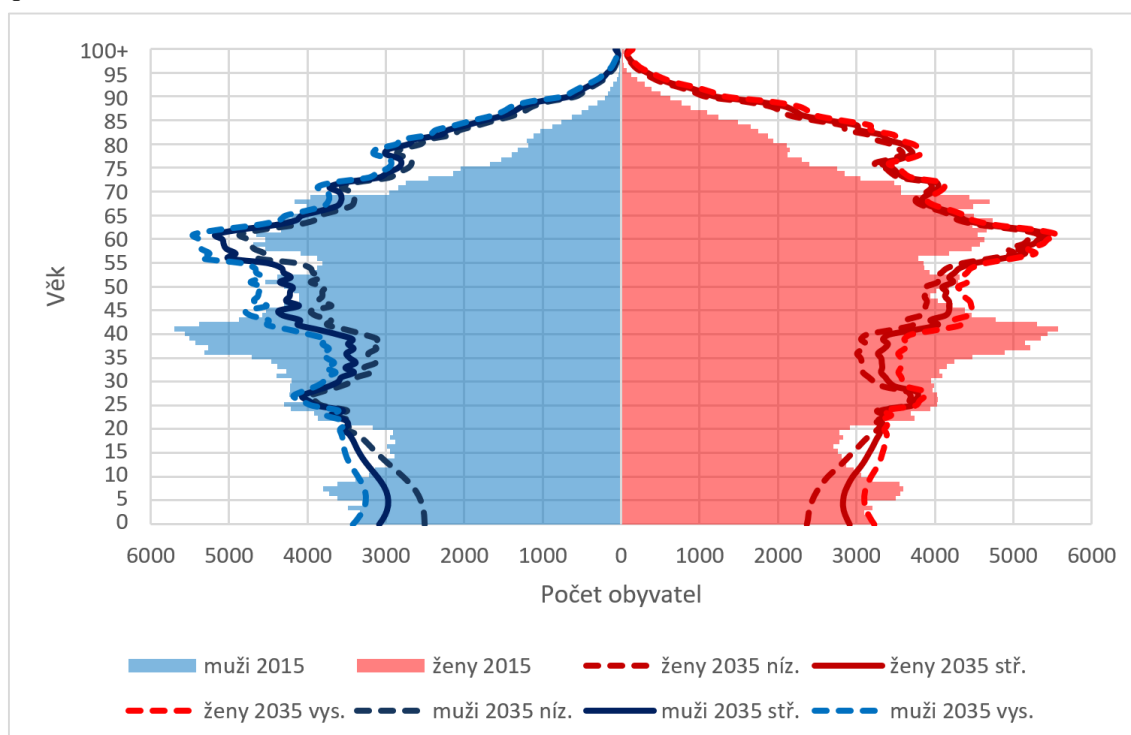


**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).



Na obrázku 57 je zobrazena prognóza věkové struktury obyvatel Jihočeského kraje k 31. 12. 2035 se spolehlivostí 95% a stejná prognóza se spolehlivostí 75% se nachází v příloze 13. Rozdíly mezi variantami jsou způsobeny především kvůli různým variantám počtu přistěhovalých v produktivním věku, tzn. u mužů 30 až 60 let a 30 až 50 let u žen. Výrazné rozdíly v nejnižších věcích jsou způsobeny odlišnou úrovní úhrnné plodnosti žen. Očekává se výrazné stárnutí obyvatel Jihočeského kraje a tato tematika může být inspirací pro další práce. Zároveň se očekává snížení počtu narozených, tato problematika může přispět ke kvalitnějšímu plánování školských a předškolních aktivit.

**Obr. 57: Porovnání prognózy věkové struktury mužů a žen k 31. 12. 2035 oproti 31. 12. 2015 s 95% spolehlivostí**



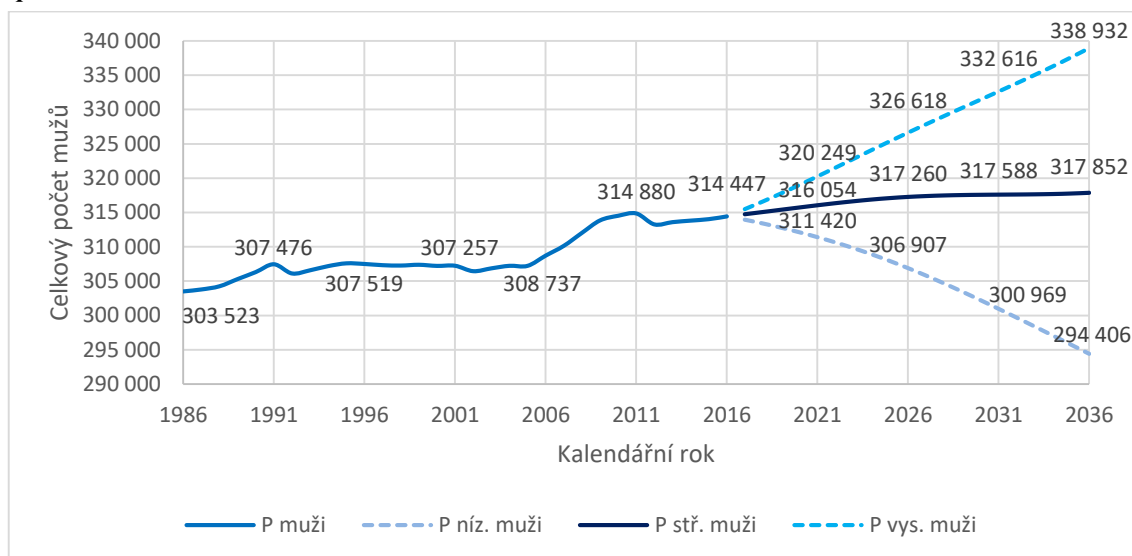
**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

Vývoj celkového počtu mužů v Jihočeském kraji je zpracován na obrázku 58 včetně jejich prognózy k 31. 12. 2035 se spolehlivostí 95% a v příloze 14 je zobrazena tato prognóza s 75% spolehlivostí. Celkový počet mužů v Jihočeském kraji má přibližně lineárně rostoucí trend, kolem kterého právě tento počet kolísá. Prognóza počtu mužů je vytvořena od 31. 12. 2016 do 31. 12. 2035. Nízká a velmi málo pravděpodobná varianta předpokládá výrazné snížení počtu mužů v Jihočeském kraji. Střední a nejpravděpodobnější varianta očekává zvýšení počtu mužů k 31. 12. mezi roky 2015 až 2035 o 3405 mužů. Tento nárůst je způsoben převážně předpokládaným počtem přistěhovalých mužů. Vysoká a málo pravděpodobná varianta očekává výrazné zvýšení počtu mužů.

Na obrázku 59 je zobrazen vývoj celkového počtu žen k 31. 12. mezi roky 1985 až 2015 v Jihočeském kraji a do konce roku 2035 je vytvořena jejich prognóza s 95% spolehlivostí. V praxi se často používají prognózy s nižší spolehlivostí, z tohoto důvodu byla vytvořena

prognóza s 75% spolehlivostí, která je zobrazena v příloze 15. Z výchozích dat je patrné, že docházelo k nárůstu počtu žen a podle střední varianty prognózy lze očekávat, že tento trend bude pokračovat. Střední varianta předpokládá zvýšení počtu žen mezi roky 2016 až 2036 o 5700 žen. Stejně jako u mužů je toto zvýšení způsobeno převážně nárůstem počtu přistěhovalých žen. Nízká varianta předpokládá výrazný pokles počtu žen, naopak vysoká varianta očekává výrazné zvýšení jejich počtu. Obě varianty jsou ale málo pravděpodobné. Porovnájí-li se počty mužů a žen, tak celkově je více žen v Jihočeském kraji. To lze přisuzovat nadúmrtnosti mužů.

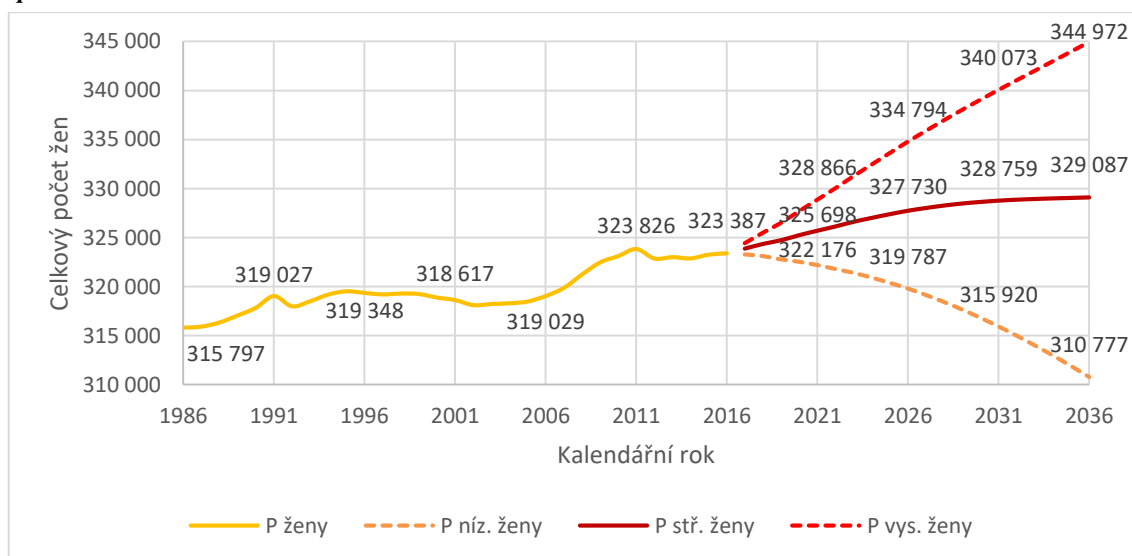
**Obr. 58:** Vývoj celkového počtu mužů k 1. 1. mezi lety 1986–2016 a jeho prognóza do roku 2035 se spolehlivostí 95 %



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** P představuje celkový počet mužů k 1. 1.

**Obr. 59:** Vývoj celkového počtu žen k 1. 1. mezi lety 1986–2016 a jeho prognóza do roku 2035 se spolehlivostí 95 %



**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** P představuje celkový počet žen k 1. 1.

## Kapitola 8

### **Komparace výsledků prognózy Jihočeského kraje s 95% spolehlivostí vytvořené z dlouhé výchozí časové řady se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu**

Tato kapitola se zaměřuje na porovnání výsledků prognózy zpracované z dlouhé výchozí časové řady se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka vystupujícího pod zkratkou VÚV TGM (Burcin a kol., 2014) a projekcí Českého statistického úřadu (ČSÚ, 2014e), pro tuto instituci se používá označení ČSÚ.<sup>5</sup> Cílem je porovnat dílčí prognózu úmrtnosti, plodnosti, migrace a celkovou populační prognózu věkové struktury obyvatel Jihočeského kraje zpracované z dlouhé výchozí časové řady s již zmíněnými odhady těchto institucí.

#### **8.1 Porovnání dílčích prognóz úmrtnosti, plodnosti a migrace se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu**

Při porovnání prognózy naděje dožití při narození v Jihočeském kraji zpracované z dlouhé výchozí časové řady (Tab. 11) se střední variantou prognózy vytvořené v rámci VÚV TGM (Burcin a kol., 2014) a projekcí ČSÚ (ČSÚ, 2014e) bylo zjištěno, že všechny hodnoty naděje dožití u obou variant vybraných pro srovnání leží v 95% intervalu spolehlivosti prognózy zpracované z dlouhé výchozí časové řady. Projekce ČSÚ očekává rychlejší nárůst naděje dožití při narození u mužů a ještě výraznější nárůst u žen oproti oběma středním variantám prognóz naděje dožití zveřejněných v tabulce 11. Od roku 2030 se předpokládá podle projekce ČSÚ nižší

---

<sup>5</sup> Projekce ČSÚ na úrovni krajů byla publikována na stránkách ČSÚ pouze v jedné variantě (ČSÚ, 2014e).

tempo nárůstu naděje dožití při narození. Střední varianta prognózy zpracovaná v rámci VÚV TGM očekává všeobecně nižší naději dožití při narození oproti střední variantě prognózy vytvořené z dlouhé výchozí časové řady a oproti projekci ČSÚ. Při porovnání odchylek mezi střední variantou prognózy naděje dožití při narození vytvořené z dlouhé výchozí časové řady a projekci ČSÚ a střední variantou prognózy zpracované v rámci VÚV TGM, bylo zjištěno, že střední varianta prognózy naděje dožití mužů zpracovaná z dlouhé výchozí časové řady se více blíží projekci ČSÚ, naopak u žen se střední varianta prognózy naděje dožití vytvořená z dlouhé výchozí časové řady více blíží střední variantě prognózy vytvořené v rámci VÚV TGM. Všechny zde představené odhady budoucího vývoje naděje dožití při narození předpokládají, že se úmrtnost mužů a žen se bude sblížovat.

**Tab. 11: Prognóza naděje dožití při narození chlapců a dívek v Jihočeském kraji se spolehlivostí 95 % zpracovaná z dlouhé výchozí časové řady včetně střední varianty prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce Českého statistického úřadu mezi lety 2016–2035**

Pohlaví	Naděje dožití	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
Muži	e(0) níz.	75,37	75,60	75,84	76,08	76,31	76,55	76,78	77,01	77,25	77,48
	e(0) stř.	76,12	76,38	76,64	76,90	77,15	77,41	77,66	77,91	78,16	78,40
	e(0) vys.	76,86	77,14	77,42	77,70	77,97	78,24	78,51	78,78	79,04	79,30
	e(0) ČSÚ 2014	76,16	76,42	76,69	76,95	77,22	77,48	77,74	78,01	78,27	78,54
	e(0) VÚV TGM 2014	76,27	76,51	76,74	76,97	77,20	77,42	77,64	77,86	78,08	78,30
Ženy	e(0) níz.	81,46	81,60	81,40	81,90	81,91	82,01	82,32	82,37	82,56	82,76
	e(0) stř.	82,04	82,17	82,03	82,60	82,61	82,76	83,09	83,15	83,37	83,59
	e(0) vys.	82,60	82,73	82,64	83,27	83,28	83,48	83,83	83,90	84,15	84,38
	e(0) ČSÚ 2014	81,91	82,14	82,37	82,61	82,84	83,07	83,31	83,54	83,77	84,00
	e(0) VÚV TGM 2014	81,76	81,96	82,16	82,35	82,54	82,73	82,92	83,10	83,27	83,43

Pohlaví	Naděje dožití	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
Muži	e(0) níz.	77,71	77,94	78,16	78,39	78,62	78,84	79,07	79,29	79,51	79,73
	e(0) stř.	78,65	78,89	79,13	79,37	79,61	79,84	80,08	80,31	80,54	80,78
	e(0) vys.	79,56	79,82	80,07	80,32	80,57	80,82	81,06	81,31	81,55	81,79
	e(0) ČSÚ 2014	78,80	79,07	79,33	79,60	79,86	80,03	80,20	80,37	80,54	80,71
	e(0) VÚV TGM 2014	78,51	78,72	78,92	79,12	79,31	79,50	79,69	79,88	80,06	80,24
Ženy	e(0) níz.	82,86	83,06	83,22	83,36	83,54	83,69	83,84	84,01	84,16	84,32
	e(0) stř.	83,71	83,93	84,10	84,26	84,45	84,61	84,78	84,96	85,12	85,29
	e(0) vys.	84,52	84,76	84,94	85,11	85,32	85,49	85,67	85,86	86,03	86,21
	e(0) ČSÚ 2014	84,24	84,47	84,70	84,94	85,17	85,31	85,45	85,59	85,73	85,87
	e(0) VÚV TGM 2014	83,59	83,75	83,91	84,06	84,21	84,36	84,51	84,65	84,79	84,93

**Zdroj:** ČSÚ (2014e), Burcin a kol. (2014), vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** e(0) představuje naději dožití při narození.

V tabulce 12 je zobrazena prognóza úhrnné plodnosti v Jihočeském kraji vytvořená z dlouhé výchozí časové řady. Tato prognóza je porovnána se střední variantou prognózy zpracované v rámci VÚV TGM (Burcin a kol., 2014) a s projekcí ČSÚ (ČSÚ, 2014e). Podle střední varianty prognózy úhrnné plodnosti zpracované z dlouhé výchozí časové řady se očekává, že úhrnná plodnost bude pokračovat ve svém současném rychlém růstu a vzroste z hodnoty 1,59 živě narozených dětí na jednu ženu v roce 2016 na hodnotu 1,74 živě narozených dětí na jednu ženu v roce 2035. Projekce ČSÚ předpokládá pouze mírný nárůst úhrnné plodnosti z hodnoty 1,51 v roce 2016 na hodnotu 1,53 v roce 2035. Jako nejrealističtější se jeví střední varianta prognózy zpracovaná v rámci VÚV TGM, která předpokládá nárůst úhrnné plodnosti z hodnoty

1,55 živě narozených dětí na jednu ženu v roce 2016 na hodnotu 1,61 živě narozených dětí na jednu ženu v roce 2035. Obecně lze říci, že střední varianta prognózy úhrnné plodnosti zpracovaná z dlouhé výchozí časové řady přepokládá větší nárůst úhrnné plodnosti než ostatní existující odhady.

**Tab. 12: Prognóza úhrnné plodnosti Jihočeského kraje se spolehlivostí 95 % zpracovaná z dlouhé výchozí časové řady včetně střední varianty prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce Českého statistického úřadu mezi lety 2016–2035**

Úhrnná plodnost	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
úp níz.	1,567	1,567	1,568	1,568	1,569	1,568	1,567	1,566	1,564	1,561
úp stř.	1,594	1,607	1,618	1,629	1,639	1,649	1,658	1,666	1,674	1,682
úp vys.	1,616	1,638	1,656	1,672	1,687	1,701	1,714	1,727	1,738	1,750
úp ČSÚ 2014	1,513	1,514	1,514	1,514	1,515	1,515	1,515	1,516	1,516	1,516
úp VÚV TGM 2014	1,548	1,553	1,560	1,565	1,571	1,577	1,581	1,585	1,589	1,592

Úhrnná plodnost	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
úp níz.	1,557	1,553	1,548	1,542	1,535	1,527	1,518	1,507	1,494	1,478
úp stř.	1,689	1,696	1,703	1,710	1,716	1,722	1,727	1,733	1,738	1,744
úp vys.	1,760	1,770	1,780	1,790	1,799	1,807	1,816	1,824	1,832	1,840
úp ČSÚ 2014	1,517	1,517	1,517	1,518	1,518	1,521	1,524	1,527	1,530	1,533
úp VÚV TGM 2014	1,594	1,597	1,599	1,601	1,603	1,606	1,608	1,610	1,612	1,614

**Zdroj:** ČSÚ (2014e), Burcin a kol. (2014), vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** úp představuje úhrnnou plodnost.

V tabulce 13 je zobrazen vývoj migračního salda Jihočeského kraje dohromady za muže a ženy, včetně prognózy zpracované z dlouhé výchozí časové řady, střední varianty prognózy vytvořené v rámci VÚV TGM (Burcin a kol., 2014) a projekce ČSÚ (ČSÚ, 2014e). Celkový vývoj migračního salda za minulá období pro obě pohlaví dohromady byl zařazen z důvodu dosavadního popisování vývoje migračního salda zvlášť za muže a zvlášť za ženy, vývoj celkového migračního salda pro obě pohlaví dohromady nebyl doposud prezentován. Ve vývoji migračního salda se objevovaly značné výkyvy. Postupně docházelo k jeho růstu a poklesu. Tyto výkyvy byly způsobeny převážně ekonomickými podmínkami v České republice. Střední varianta prognózy migračního salda zpracovaná z dlouhé výchozí časové řady pro muže a ženy předpokládala růst tohoto salda na úroveň, jaká byla před krizí, tzn. v roce 2008. Tato skutečnost přepokládá, že do roku 2035 nenastane ekonomická krize. Při porovnání střední varianty prognózy migračního salda zpracované z dlouhé výchozí časové řady a projekce ČSÚ bylo zjištěno, že oba odhady migračního salda předpokládají jeho nárůst do roku 2035. U střední varianty prognózy z dlouhé výchozí časové řady se předpokládá výraznější růst migračního salda než přepokládá projekce ČSÚ. Střední varianta prognózy vytvořená v rámci VÚV TGM očekává nárůst migračního salda do roku 2026, poté tato varianta předpokládá pokles migračního salda do roku 2030 a následně do roku 2035 se předpokládá konstantní migrační saldo.

**Tab. 13: Prognóza celkového migračního salda obyvatel Jihočeského kraje se spolehlivostí 95 % zpracovaná z dlouhé výchozí časové řady včetně střední varianty prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce Českého statistického úřadu mezi lety 2016–2035, vývoj celkového migračního salda mezi lety 1993–2015**

Migrační saldo	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
MS	–	–	–	–	–	–	–	738	791	775

Migrační saldo	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
MS	677	1 004	878	412	482	94	1 024	1 229	670	2 316

Migrační saldo	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
MS	2 038	2 582	2 256	755	546	362	322	326	584	867

Migrační saldo	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
MS níž.	–22	19	59	100	143	189	237	284	329	369
MS stř.	857	947	1 030	1 110	1 188	1 264	1 337	1 406	1 471	1 528
MS vys.	1 649	1 792	1 923	2 047	2 165	2 277	2 382	2 481	2 573	2 657
MS ČSÚ 2014	552	589	629	671	717	765	812	858	902	942
MS VÚV TGM 2014	533	600	666	733	800	860	920	980	1 040	1 100

Migrační saldo	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035
MS níž.	401	427	447	462	469	468	463	452	437	422
MS stř.	1 578	1 620	1 656	1 686	1 709	1 724	1 734	1 739	1 740	1 741
MS vys.	2 731	2 798	2 858	2 911	2 958	2 998	3 031	3 061	3 087	3 113
MS ČSÚ 2014	976	1 005	1 029	1 047	1 061	1 070	1 075	1 078	1 080	1 083
MS VÚV TGM 2014	1 020	940	860	780	700	700	700	700	700	700

**Zdroj:** ČSÚ (2014e), (Burcin a kol., 2014), vlastní zpracování z dat poskytnutých ČSÚ (2017a).

**Pozn.:** MS představuje migrační saldo.

## 8.2 Porovnání očekávaného vývoje věkové struktury obyvatel Jihočeského kraje se střední variantou prognózy vytvořené v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekcí Českého statistického úřadu

Cílem této podkapitoly je porovnat věkovou strukturu prognózy obyvatel Jihočeského kraje zpracovanou z dlouhé výchozí časové řady se střední variantou prognózy vytvořené v rámci VÚV TGM (Burcin a kol., 2014) a projekcí ČSÚ (ČSÚ, 2014e). Při pozorování očekávané věkové struktury projekce ČSÚ a střední varianty prognózy vytvořené v rámci VÚV TGM lze pozorovat v tabulce 14 dva základní trendy. Prvním trendem je snižování počtu osob v nejnižších věkových skupinách, které očekává prohlubování během let. Český statistický úřad předpokládá výraznější snížení počtu osob v těchto věkových skupinách než střední varianta prognózy struktury obyvatel vytvořená v rámci VÚV TGM. Druhým trendem je zvyšování počtu osob v nejvyšších věkových skupinách. Například ve věkové skupině 90 a více let se počet mužů podle projekce ČSÚ mezi 1. 1. 2017 a 1. 1. 2036 téměř zčtyřnásobí a u žen se jejich počet v tomto období téměř ztrojnásobí. Obecně se očekává více mužů v nízkých věkových skupinách, naopak ve vyšších věcích se předpokládá větší zastoupení žen než mužů. Podle projekce ČSÚ i podle střední varianty prognózy zpracované v rámci VÚV TGM se očekává celkové snížení počtu mužů i žen

v Jihočeském kraji. Střední varianta prognózy vytvořená v rámci VÚV TGM předpokládá k 1. 1. 2036 výraznější snížení počtu obyvatel než projekce ČSÚ.

**Tab. 14: Střední varianta prognózy věkové struktury mužů a žen v Jihočeském kraji k 1. 1. vytvořená v rámci Výzkumného ústavu vodohospodářského T. G. Masaryka a projekce věkové struktury mužů a žen v Jihočeském kraji k 1. 1. zpracovaná Českým statistickým úřadem, roky 2017, 2026 a 2036**

Muži						
Věk	Projekce ČSÚ 2014			Střední varianta prognózy VÚV TGM 2014		
	2017	2026	2036	2017	2026	2036
0–4	16 533	13 969	12 841	16 706	14 647	13 455
5–9	18 029	15 580	13 231	18 064	16 109	13 928
10–14	15 709	16 857	14 368	15 749	17 018	15 010
15–19	14 655	18 140	15 746	14 690	18 239	16 243
20–24	17 427	15 502	17 083	17 445	15 662	17 239
25–29	20 823	15 072	18 296	20 795	15 226	18 255
30–34	21 523	18 667	15 920	21 437	18 670	15 888
35–39	25 233	20 848	15 764	25 221	20 885	15 886
40–44	26 789	21 768	18 965	26 787	21 817	18 973
45–49	21 450	25 402	20 539	21 480	25 379	20 493
50–54	21 136	24 825	21 013	21 112	24 757	20 940
55–59	19 985	20 072	24 110	19 952	20 022	23 864
60–64	22 353	19 512	23 353	22 329	19 389	22 958
65–69	20 239	18 699	18 615	20 273	18 522	18 184
70–74	14 586	18 451	17 162	14 571	18 287	16 594
75–79	8 369	15 072	14 917	8 394	14 961	14 331
80–84	5 432	8 181	12 655	5 495	8 147	12 152
85–89	2 797	3 612	7 854	2 856	3 656	7 651
90+	851	1 683	3 186	913	1 841	3 530
Celkem	313 919	311 912	305 618	314 269	313 233	305 573

Ženy						
Věk	Projekce ČSÚ 2014			Střední varianta prognózy VÚV TGM 2014		
	2017	2026	2036	2017	2026	2036
0–4	15 540	13 155	12 090	15 668	13 808	12 698
5–9	17 265	14 676	12 455	17 219	15 141	13 106
10–14	14 940	15 821	13 526	14 916	15 816	14 049
15–19	13 902	17 402	14 839	13 882	17 286	15 139
20–24	16 574	14 936	16 154	16 580	14 979	16 068
25–29	20 005	14 569	17 740	19 932	14 652	17 452
30–34	20 463	17 832	15 541	20 350	17 780	15 353
35–39	24 129	20 257	15 425	24 078	20 195	15 445
40–44	26 397	21 010	18 409	26 350	20 966	18 343
45–49	21 040	25 025	20 379	21 035	24 889	20 195
50–54	20 744	25 224	20 950	20 674	25 088	20 777
55–59	20 124	20 409	24 718	20 094	20 337	24 405
60–64	22 869	20 260	24 792	22 816	20 194	24 510
65–69	22 521	20 073	19 919	22 481	20 051	19 760
70–74	17 827	21 132	19 028	17 767	20 984	18 745
75–79	12 087	18 964	17 627	12 054	18 741	17 247
80–84	9 307	12 249	16 649	9 302	12 025	15 929
85–89	5 673	6 644	12 003	5 706	6 519	11 241
90+	2 379	4 088	6 911	2 512	3 949	6 343
Celkem	323 786	323 726	319 155	323 417	323 402	316 806

**Zdroj:** ČSÚ (2014e), (Burcin a kol., 2014), vlastní zpracování.

V tabulce 15 je prezentována prognóza věkové struktury mužů a žen Jihočeského kraje k 1. 1. ve vybraných letech vytvořená z dlouhé výchozí časové řady s 95% spolehlivostí. Nízká varianta předpokládá výrazný pokles počtu obyvatel k 1. 1. roku 2017, 2026 a 2036. Tato varianta

je ale málo pravděpodobná, přibližně 2,5 %. Naopak vysoká varianta očekává výrazný nárůst počtu obyvatel do roku 2035. Pravděpodobnost této varianty se pohybuje přibližně na úrovni 2,5 %. Jako nejpravděpodobnější se jeví střední varianta, která očekává mírný nárůst počtu obyvatel.

**Tab. 15: Prognóza věkové struktury mužů a žen Jihočeského kraje zpracovaná z dlouhé výchozí časové řady k 1. 1. se spolehlivostí 95 %, roky 2017, 2026 a 2036**

Muži									
Věk	Nízká varianta prognózy			Střední varianta prognózy			Vysoká varianta prognózy		
	2017	2026	2036	2017	2026	2036	2017	2026	2036
0–4	16 828	14 274	12 581	16 909	15 763	15 100	16 984	16 861	16 676
5–9	18 067	15 937	13 124	18 071	16 712	15 033	18 074	17 344	16 379
10–14	15 757	16 973	14 559	15 763	17 036	16 029	15 768	17 098	17 139
15–19	14 697	18 283	16 171	14 703	18 361	17 007	14 710	18 439	17 713
20–24	17 487	15 896	17 559	17 501	16 025	17 799	17 516	16 159	18 050
25–29	20 987	15 782	19 073	21 045	16 094	19 596	21 103	16 423	20 153
30–34	21 369	18 720	16 402	21 458	19 475	17 489	21 541	20 242	18 651
35–39	25 069	20 248	15 819	25 145	21 208	17 289	25 215	22 159	18 861
40–44	26 847	20 826	18 133	26 939	21 752	20 017	27 018	22 628	21 948
45–49	21 491	24 516	19 107	21 564	25 555	21 200	21 624	26 467	23 236
50–54	21 173	24 239	19 464	21 225	25 159	21 399	21 270	25 950	23 207
55–59	20 006	19 771	22 712	20 041	20 349	24 538	20 072	20 864	26 182
60–64	22 360	19 186	22 274	22 403	19 607	23 717	22 442	19 996	25 016
65–69	20 234	18 271	17 735	20 272	18 649	18 681	20 307	19 003	19 559
70–74	14 444	17 914	16 185	14 483	18 313	16 926	14 519	18 688	17 625
75–79	8 287	14 261	13 746	8 321	14 699	14 461	8 353	15 115	15 144
80–84	5 302	7 335	11 066	5 334	7 675	11 842	5 365	8 004	12 598
85–89	2 737	3 060	6 298	2 759	3 267	6 946	2 781	3 472	7 599
90+	819	1 418	2 396	827	1 558	2 782	836	1 704	3 197
Celkem	313 960	306 907	294 406	314 764	317 260	317 852	315 497	326 618	338 932

Ženy									
Věk	Nízká varianta prognózy			Střední varianta prognózy			Vysoká varianta prognózy		
	2017	2026	2036	2017	2026	2036	2017	2026	2036
0–4	15 791	13 521	11 937	15 873	14 925	14 311	15 944	15 961	15 799
5–9	17 323	15 121	12 506	17 328	15 862	14 310	17 332	16 466	15 588
10–14	14 967	15 912	13 795	14 972	15 990	15 196	14 978	16 064	16 260
15–19	13 956	17 527	15 316	13 963	17 611	16 122	13 970	17 692	16 798
20–24	16 616	15 207	16 436	16 628	15 310	16 649	16 639	15 410	16 857
25–29	19 929	14 915	18 009	20 000	15 189	18 484	20 066	15 453	18 941
30–34	20 177	17 642	15 620	20 266	18 463	16 742	20 345	19 234	17 824
35–39	23 932	19 544	15 321	23 985	20 451	16 674	24 033	21 291	17 992
40–44	26 378	20 326	17 782	26 416	20 905	19 292	26 450	21 434	20 742
45–49	21 031	24 545	19 424	21 055	24 991	20 809	21 076	25 393	22 111
50–54	20 713	25 041	20 167	20 729	25 349	21 079	20 744	25 629	21 925
55–59	20 112	20 351	24 130	20 123	20 533	24 831	20 133	20 701	25 471
60–64	22 775	20 157	24 484	22 792	20 327	25 031	22 808	20 489	25 542
65–69	22 482	19 891	19 694	22 503	20 078	20 096	22 522	20 254	20 476
70–74	17 804	20 932	18 775	17 832	21 189	19 188	17 857	21 424	19 570
75–79	12 024	18 599	17 190	12 057	18 985	17 738	12 088	19 340	18 241
80–84	9 198	11 568	15 534	9 239	11 978	16 338	9 277	12 364	17 086
85–89	5 613	5 853	10 137	5 649	6 184	11 015	5 683	6 506	11 869
90+	2 445	3 137	4 520	2 466	3 410	5 182	2 485	3 690	5 880
Celkem	323 266	319 787	310 777	323 876	327 730	329 087	324 431	334 794	344 972

**Zdroj:** vlastní zpracování z dat poskytnutých ČSÚ (2017a).



Při porovnání střední varianty prognózy struktury obyvatel Jihočeského kraje vytvořené z dlouhé výchozí časové řady (Tab. 15) se střední variantou prognózy vytvořené v rámci VÚV TGM (Burcin a kol., 2014) a projekcí ČSÚ (ČSÚ, 2014e), obě jsou zobrazeny v tabulce 14, je patrný jeden rozdíl. Tímto rozdílem je očekávaná změna počtu obyvatel Jihočeského kraje mezi 1. 1. 2017 a 1. 1. 2035, kdy při porovnání celkového vývoje počtu obyvatel ve střední variantě prognózy vytvořené z dlouhé výchozí časové řady se očekává nárůst počtu obyvatel, zatímco zmíněné odhady očekávaného vývoje v tabulce 14 přepokládají celkový pokles počtu obyvatel. Tato skutečnost je zapříčiněna odlišným zpracováním středních variant dílčích prognóz, případně odlišných parametrů projekce ČSÚ. Ve střední variantě prognózy zpracované z dlouhé výchozí časové řady je oproti střední variantě prognózy vytvořené v rámci VÚV TGM a projekci ČSÚ výrazně vyšší očekávaný počet přistěhovalých a také vyšší předpokládaný počet narozených. Tento nárůst způsobuje i mírný růst celkového počtu obyvatel v Jihočeském kraji v této variantě prognózy. Stejně jako u projekce ČSÚ a střední variantě prognózy vytvořené v rámci VÚV TGM jsou i ve střední variantě prognózy zpracované z dlouhé výchozí časové řady vidět stejné trendy, a to ubývání počtu osob v nejnižších věcích a nárůst počtu osob v nejvyšších věcích. Pokles počtu osob v nejnižších věcích se očekává pomalejší než ve zmíněných odhadech očekávaného vývoje struktury obyvatel v tabulce 14. Naopak nižší nárůst v nejvyšších věcích předpokládá střední varianta prognózy zpracovaná z dlouhé výchozí časové řady. Tato střední varianta předpokládá rovněž vyšší počet osob v produktivním věku právě kvůli očekávanému vyššímu počtu přistěhovalých. Společným charakterem obou středních variant prognóz a projekce ČSÚ je obecně vyšší počet mužů v nižších věkových skupinách a naopak převažující počet žen ve vyšších věkových skupinách.

## Kapitola 9

### Závěr

Základním cílem této diplomové práce bylo představit vybrané přístupy k tvorbě populačních prognóz pro regionální úroveň s uvažováním různé délky časové řady a různé kvality dat. Jednotlivé postupy byly prakticky aplikovány za účelem vytvoření populační prognózy Jihočeského kraje. Pro tento účel byly stanoveny čtyři cíle.

Prvním cílem bylo popsat vybrané metody pro tvorbu populačních prognóz. Při pozorování současných trendů v této oblasti bylo zjištěno, že mezi nejpoužívanější metody pro zhotovení populačních prognóz jednoznačně patří kohortně komponentní metoda bez migrace a kohortně komponentní metoda s migrací. Obě metody byly v této práci podrobně matematicky popsány. K dílčímu cíli patřila charakteristika vybraných demografických modelů pro tvorbu dílčí prognózy úmrtnosti, plodnosti a migrace. Obecně bylo zjištěno, že se pro odhad těchto dílčích prognóz používá několik přístupů, mezi které patří odborný odhad, analytický přístup prostřednictvím demografických modelů, vztahů a funkcí nebo jejich kombinace. Pro tuto diplomovou práci byl zvolen analytický přístup pro vytvoření dílčích prognóz. Současným trendem je používání Lee-Carterových modelů pro tyto dílčí prognózy. Existuje i řada dalších modelů, které se zpravidla používají k vyrovnání intenzit úmrtnosti a plodnosti, ale mohou být použity pro účely dílčích prognóz. Z těchto modelů byl pro účely dílčí prognózy úmrtnosti popsán Heligman-Pollardův model a pro dílčí prognózu plodnosti byla charakterizována Beta funkce. Vzhledem ke skutečnosti, že data za migraci nejsou často k dispozici, byla popsána nepřímá metoda výpočtu migračního salda. Pro extrapolaci parametrů byly stručně charakterizovány základní principy analýzy časových řad.

Druhým cílem bylo zpracování prognózy Jihočeského kraje z krátké výchozí časové řady. Tato časová řada byla zpracována z méně kvalitních a špatně dostupných dat, aby simulovala údaje z velmi malých územních celků nebo data za rozvojové země. Před zpracováním celé prognózy bylo nejprve nutné zpracovat dílčí prognózu úmrtnosti, plodnosti a migrace. Pro dílčí prognózu úmrtnosti byl zvolen osmiparametrový Heligman-Pollardův model a pro dílčí prognózu plodnosti byla vybrána Beta funkce. Problémem, který bylo nutné vyřešit, byly vzájemné vztahy mezi odhadnutými parametry funkcí, kvůli kterým nebylo možné modelovat každý parametr zvlášť prostřednictvím jejich extrapolace. V této práci byl navržen specifický způsob řešení této problematiky prostřednictvím faktorové analýzy, kdy se jejím prostřednictvím identifikovaly

vzájemné vazby mezi parametry. Byla vypočítána faktorová skóre pro jednotlivé faktory, které byly extrapolovány na základě trendových funkcí. Za předpokladu lineárního vztahu mezi parametry dané funkce a příslušnými faktorovými skóry byly zpětně domodelovány prostřednictvím jednoduché lineární regrese původní parametry. Při jiném než lineárním vztahu mezi parametry a faktorovými skóry by lineární regrese musela být modifikována. Tento poměrně jednoduchý způsob může prosadit častější používání složitějších modelů a funkcí pro zpracování nejen demografických, ale například i ekonomických prognóz. V této práci se podařilo nalézt jedno omezení tohoto přístupu, pro který jsou zapotřebí kvalitní odhady parametrů daných modelů a funkcí. Toto omezení se projevilo v této práci při zpracovávání dílčí prognózy úmrtnosti, kdy odhady parametrů Heligman-Pollardova modelu z krátké výchozí časové řady byly málo kvalitní a z tohoto důvodu byla celá dílčí prognóza vychýlena. Tím může vzniknout prostor pro další práce na zpracování programu pro odhad parametrů Heligman-Pollardova modelu z kvocientů úmrtnosti podle jednotek věku.

Do Beta funkce vstupovaly parametry jako průměrný věk matek, rozptyl průměrného věku matek a úhrnná plodnost. Následně bylo nutné vypočítat migrační saldo. Protože Český statistický úřad nepublikuje počty přistěhovalých, počty vystěhovalých ani migrační saldo za kraje podle jednotek věku, bylo nutné tato data odhadnout. Pro tento účel byl zvolen odhad prostřednictvím nepřímé metody výpočtu migračního salda. Dílčí prognóza migrace byla zpracována z 25%, 50% a 75% kvartilu migračního salda podle jednotek věku. Celá prognóza z krátké výchozí časové řady byla zpracována kohortně komponentní metodou s postupným zakomponováním migračního salda. Vzhledem k ne zcela kvalitním odhadům parametrů Heligman-Pollardova modelu u dílčí prognózy úmrtnosti a také kvůli menší kvalitě dat, nebyla tato populační prognóza v závěru porovnána s ostatními prognózami.

Třetí cíl směřoval k vytvoření kvalitní populační prognózy Jihočeského kraje z dlouhé výchozí časové řady. Pro tvorbu této populační prognózy byla zvolena kohortně komponentní metoda s migrací s dvoufázovým zahrnutím imigrace a emigrace do modelu. Před vytvořením této prognózy je nutné nejprve zpracovat dílčí prognózy úmrtnosti, plodnosti, imigrace a emigrace. Pro dílčí prognózu úmrtnosti byl zvolen Lee-Carterův model úmrtnosti a následně jeho index úrovně úmrtnosti byl prognózován prostřednictvím ARIMA modelu. Prognóza plodnosti byla zpracována Lee-Carterovým modelem plodnosti a index úrovně plodnosti byl následně prognózován rovněž prostřednictvím ARIMA modelu. Pro dílčí prognózu imigrace byl zvolen Lee-Carterův model imigrace, pro dílčí prognózu emigrace byl vybrán Lee-Carterův model emigrace. Vzhledem ke skutečnosti, že dlouhá výchozí časová řada indexu úrovně imigrace a indexu úrovně emigrace nedosahovala minimální hranice pro použití ARIMA modelů, byl pro jejich odhad zvolen logaritmický trend.

Čtvrtým a posledním cílem bylo vzájemně porovnat prognózu věkové struktury obyvatel Jihočeského kraje vytvořenou z dlouhé výchozí časové řady se střední variantou prognózy vytvořené v rámci VÚV TGM a s projekcí ČSÚ. Nejprve byly porovnány dílčí prognózy úmrtnosti, plodnosti a migrace. Při porovnání obou středních variant prognóz naděje dožití při narození a projekce ČSÚ bylo zjištěno, že jak střední varianta prognózy vytvořená v rámci VÚV TGM, tak i projekce ČSÚ se pohybují ve stanovených mezích prognózy naděje dožití při

narození zpracované z dlouhé výchozí časové řady. Při porovnání střední varianty dílčí prognózy plodnosti zpracované z dlouhé výchozí časové řady bylo zjištěno, že je optimističtější oproti projekci ČSÚ a oproti střední variantě prognózy vytvořené v rámci VÚV TGM. Střední varianta dílčí prognózy migrace vytvořená z dlouhé výchozí časové řady očekává vyšší přírůstek stěhování než oba zveřejněné odhady budoucího vývoje v rámci VÚV TGM a ČSÚ. Tím je poté ovlivněna i očekávaná struktura počtu obyvatel, která ve střední variantě prognózy vytvořené z dlouhé výchozí časové řady očekává vyšší počty osob v nejnižších věkových skupinách a v produktivním věku. Zároveň také očekává k 1. 1. 2036 vyšší přírůstek celkového počtu obyvatel oproti projekci ČSÚ a oproti střední variantě prognózy vytvořené v rámci VÚV TGM, které obě předpokládají pokles celkového počtu obyvatel do tohoto roku.

Tato diplomová práce má široké možnosti uplatnění pro další výzkum. Může sloužit jako návod pro vytvoření prognózy pro další práce nebo instituce. Vzhledem ke skutečnosti, že zde byly popsány pouze vybrané přístupy k tvorbě populační prognózy a k vytvoření dílčích prognóz úmrtnosti, plodnosti a migrace, může být tato práce dále rozšířena o další metody. Bylo zde nastíněno několik problémů s odhadem parametrů Heligman-Pollardova modelu. Řešení těchto problémů by mohlo být inspirací pro další práce. Velký potenciál k využití má také navržené uplatnění faktorové analýzy k identifikaci vztahů mezi parametry, a to nejen u demografických, ale i matematických a ekonomických modelů a funkcí. Populační prognóza věkové struktury Jihočeského kraje zpracovaná z dlouhé výchozí časové řady může sloužit pro řadu odvozených prognóz například k odhadu počtu žáků v Jihočeském kraji nebo k řešení otázky kapacity domovů pro seniory v Jihočeském kraji.

## SEZNAM POUŽITÉ LITERATURY

- ARLT, Josef a Markéta ARLTOVÁ. 2003. *Finanční časové řady: [vlastnosti, metody modelování, příklady a aplikace]*. Praha: Grada. ISBN 80-247-0330-0.
- ARLTOVÁ, Markéta a Josef ARLT. 2009. *Grafické metody analýzy ekonomických časových řad*. Statistika, 32.11: 483–493. ISSN 0322-788x.
- ARRIAGA, Eduardo E. 1984. *Measuring and explaining the change in life expectancies*. Demography, 21.1: 83–96.
- AUBRECHT, Vratislav. 2016. *Globální trendy ve vývoji úmrtnosti na HIV/AIDS*. Diplomová práce. Praha: Univerzita Karlova v Praze, Přírodovědecká fakulta. Praha. Vedoucí práce: RNDr. Olga Kurtinová, Ph. D.
- BEN-DAYA, Mohamed, Salih O. Duffuaa, Abdul Raouf, Jezdimir Knezevic, Daoud Ait-Kadi, EDITORS. 2009. *Handbook of maintenance management and engineering*. New York: Springer. ISBN 978-1-84882-471-3.
- BIRCH, Leann L., J. O. FISHER, K. GRIMM-THOMAS, C. N. MARKEY, R. SAWYER a S. L. JOHNSON. 2001. *Confirmatory factor analysis of the Child Feeding Questionnaire: a measure of parental attitudes, beliefs and practices about child feeding and obesity proneness*. Appetite, 36.3: 201–210.
- BOWLEY, A. L. 1924. *Births and population in Great Britain*. The Economic Journal 34, 188–192.
- BOX, George E. P. a David A. PIERCE. 1970. *Distribution of residual autocorrelations in autoregressive–integrated moving average time series models*. Journal of the American statistical Association, 65.332: 1509–1526.
- BROUHNS, Natacha, Michel DENUIT a Jeroen K. A VERMUNT. 2002. *Poisson log–bilinear regression approach to the construction of projected lifetables*. Insurance: Mathematics and Economics, 31.3: 373–393.
- BURCIN, Boris, Klára TESÁRKOVÁ a Luděk ŠÍDLO. 2010. *Nejpoužívanější metody vyrovnávání a extrapolace křivky úmrtnosti a jejich aplikace na českou populaci*. Demografie. Roč. 52, č. 2, str. 77–89. ISSN 0011-8265.
- BURCIN, Boris a Tomáš KUČERA. 2011. *Prognóza vývoje obyvatelstva České republiky do roku 2070*. In: Bartoňová, Dagmar a kol.: *Demografická situace České republiky: proměny a kontexty 1993–2008*. Praha: Sociologické nakladatelství SLON, 2010, s. 181–212. ISBN 978-80-7419-024-7.

- BURCIN, Boris a Tomáš KUČERA. 2004. *Nová kmenová prognóza populačního vývoje České republiky (2003–2065)*. Demografie. Roč. 46, č. 2, str. 100–111. ISSN 0011-8265.
- BURCIN, Boris, Tomáš KUČERA a Luděk ŠÍDLA. 2007. *Populační prognózy a projekce – demografické okno do budoucnosti*. Geografické rozhledy, 17.1: 22–23.
- CANNAN, Edwin. 1895. *The Probability of a Cessation of the Growth of Population in England and Wales During the Next Century*. In SMITH, David P., Nathan KEYFITZ, Kenneth W. WACHTER a Hervé LE BRAS. *Mathematical demography: selected papers*. Berlin: Springer, 2013. ISBN 9783642358579.
- CARMICHAEL, Gordon A. 2016. *Fundamentals of demographic analysis: concepts : measures and methods*. Cham: Springer. Springer series on demographic methods and population analysis. ISBN 978-3-319-23254-6.
- CASELLI, Graziella. 2006. *Mortality Forecasts, Hypotheses and Methods*. In CASELLI, Graziella, Jacques VALLIN a Guillaume J. WUNSCH. 2006. *Demography: analysis and synthesis, Volume 3*. London: Elsevier, s. 177–188. ISBN 0-12-765663-4.
- CIPRA, Tomáš. 1986. *Analýza časových řad s aplikacemi v ekonomii*. Praha: Státní nakladatelství technické literatury.
- ČESKO. 2011. Ústavní zákon č. 135/2011 Sb.: *Ústavní zákon, kterým se mění ústavní zákon č. 347/1997 Sb., o vytvoření vyšších územních samosprávných celků a o změně ústavního zákona České národní rady č. 1/1993 Sb., Ústava České republiky, ve znění ústavního zákona č. 176/2001 Sb.* In: *Sbírka zákonů*. 2011, ročník 2011, částka 52, číslo 135, s. 1314. ISSN 1211-1244. Dostupné také z: [aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=z&id=23366](http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=z&id=23366)
- ČESKO. 2001. *Ústavní zákon č. 176/2001: Ústavní zákon ze dne 16. května 2001, kterým se mění ústavní zákon č. 347/1997 Sb., o vytvoření vyšších územních samosprávných celků a o změně ústavního zákona České národní rady č. 1/1993 Sb., Ústava České republiky.* In: *Sbírka zákonů*. 2001, ročník 2001, částka 67, číslo 176, s. 3674. ISSN 1211-1244. Dostupné také z: [aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=3645](http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=3645)
- ČESKO. 1997. *Ústavní zákon č. 347/1997: Ústavní zákon ze dne 3. prosince 1997 o vytvoření vyšších územních samosprávných celků a o změně ústavního zákona České národní rady č. 1/1993 Sb., Ústava České republiky.* In: *Sbírka zákonů*. 1997, ročník 1997, částka 114, číslo 347, s. 7018. ISSN 1211-1244. Dostupné také z: [aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=3105](http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=3105)
- ČESKO. 1992. *Ústava České republiky 1/1993*. In: *Sbírka zákonů*. 1992, ročník 1993, částka 1, ústavní zákon č. 1, s. 2–16. ISSN 1211-1244. Dostupné také z: [aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=z&id=22427](http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=z&id=22427)
- ČESKOSLOVENSKO. 1960. *Zákon č. 36/1960: Zákon ze dne 9. dubna 1960 o územním členění státu*. In: *Sbírka zákonů Československé socialistické republiky*. 1960, ročník 1960, částka 15, číslo 36, s. 98–100. ISSN 0322-8037. Dostupné také z: [aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=965](http://aplikace.mvcr.cz/sbirka-zakonu/ViewFile.aspx?type=c&id=965)

- D'AMATO, Valaria, RUSSOLILLO Maria. 2010. *Lee-Carter error matrix simulation: heteroschedasticity impact on actuarial valuations*. In: CORAZZA, Marco a Claudio PIZZI. *Mathematical and statistical methods for actuarial sciences and finance*. New York: Springer Verlag. ISBN 978-88-470-1480-0.
- DAS GUPTA, Monica, Jiang ZHENGHUA, Li BOHUA, Xie ZHENMING, Woojin CHUNG a Bae HWA-OK. 2003. *Why is Son Preference so Persistent in East and South Asia: a Cross-Country Study of China, India and the Republic of Korea*. *Journal of Development Studies*, vol. 40, no. 2, December, pp. 153–187.
- DIEKMANN, Andreas and Peter MITTER. 1984. *Stochastic modelling of social processes*. Orlando: Academic Press. ISBN 0122154908.
- DURBIN, James. 1960. *The fitting of time-series models*. *Revue de l'Institut International de Statistique*, Vol. 28, No. 3, pp. 233–244.
- FÁBRY, Jan. 2010. *Matematické modelování*. Praha: Oeconomica. ISBN 978-80-245-1266-2.
- FITZMAURICE, Garrett M., Nan M. LAIRD a James H. WARE. 2004. *Applied longitudinal analysis*. Hoboken: Wiley-Interscience. *Wiley series in probability and statistics*. ISBN 0-471-21487-6.
- GARCÍA-GUERRERO, Víctor Manuel. 2016. *A probabilistic method to forecast the international migration of Mexico by age and sex*. *Papeles de Población*, No 88.
- GAVRILOVA, Natalia S. a Leonid A. GAVRILOV. 2011. *Stárnutí a dlouhověkost: Zákony a prognózy úmrtnosti pro stárnoucí populace*. *Demografie*. Roč. 53, č. 2, str. 109–126. ISSN 0011-8265.
- HABARTOVÁ, Pavlína. 2009. *Projekce vývoje úmrtnosti na nemoci oběhové soustavy v České republice s využitím metody Lee-Carter*. *Demografie*. Roč. 51, č. 3, str. 173–181. ISSN 0011-8265.
- HASSANI, Sadri. 2009. *Mathematical methods: for students of physics and related fields*. 2nd ed. New York, NY: Springer. ISBN 978-0-387-09503-5.
- HELIGMAN, Larry and POLLARD, John H. 1980. *The age pattern of mortality*. *Journal of the Institute of Actuaries*, 107.01: 49–80.
- HINDE, Andrew. 1998. *Demographic methods*. New York: Arnold. ISBN 978-0-340-718926.
- HOQUE, Nazrul, Mary A. MCGEHEE, Benjamin S. BRADSHAW a EDITORS. 2013. *Applied demography and public health*. New York: Springer. ISBN 978-94-007-6139-1.
- HOWARD Anton. 2010. *Elementary linear algebra*. 10th ed. Drexel University. United States: Wiley. ISBN 9780470559918.
- HULÍK, Vladimír, Luděk ŠÍDLA a Klára TESÁRKOVÁ. 2011. *Míra účasti dětí na předškolním vzdělávání a faktory ovlivňující její regionální diferenciaci*. *Studia paedagogica*, 13.1: 13–34.
- CHAUDHRY, M. Aslam, Asghar QADIR, M. RAFIQUE a S. M. ZUBAIR. 1997. *Extension of Euler's beta function*. *Journal of Computational and Applied Mathematics*, 78(1), 19–32.

- CHUI, Charles K. a Qingtang JIANG. 2013. *Applied mathematics: Data Compression, Spectral Methods, Fourier Analysis, Wavelets and Applications*. Paris: Atlantis Press. ISBN 978-94-6239-008-9.
- JOHNSON, Kenneth M. a Richelle L. WINKLER. 2015. *Migration signatures across the decades: Net migration by age in US counties, 1950–2010*. Demographic Research, 32: 1065–1080.
- JOSHI, A. W. 1995. *Matrices and tensors in physics*. 3rd ed. New York: Wiley. ISBN 81-224-0563-0.
- KAČEROVÁ, Eva. 2004. *Mezinárodní srovnání sezónnosti sňatečnosti*. Demografie. Roč. 46, č. 3, str. 186–198. ISSN 0011-8265.
- KADERÁBKOVÁ, Jaroslava a Jitka PEKOVÁ. 2012. *Územní samospráva – udržitelný rozvoj a finance*. 1. vydání. Praha: Wolters Kluwer Česká republika, 300 s. ISBN 978-80-7357-910-4.
- KAISER, Henry F. 1974. *An index of factorial simplicity*. Psychometrika, 39.1: 31–36.
- KEILMAN, Nico, Dinh Quang PHAM a Arve HETLAND. 2001. *Norway's uncertain demographic future*. Oslo: Statistics Norway. Sosiale og økonomiske studier, 105. ISBN 82-537-5002-1.
- KING, Russell a John CONNELL. 1999. *Small worlds, global lives: islands and migration*. New York: Pinter. Island studies series (London, England). ISBN 185567548X.
- KLOSTERMAN, Richard E. 1990. *Community analysis and planning techniques*. Published by: Rowman & Littlefield. ISBN 0-8476-7651-X.
- KOCOURKOVÁ, Jiřina. 2008. *Současný „BABY-BOOM“ v České republice a rodinná politika*. Demografie. Roč. 50, č. 4., str. 240–249. ISSN 0011-8265.
- KOCOURKOVÁ, Jiřina. 2007. *Populační klima a rodinná politika*. In: Fialová, Ludmila (ed.): *Populační vývoj České republiky 2001–2006*. Praha: KDGD PřF UK, 2007. 114 s. ISBN 978-80-86561-77-6.
- KOISSI, Marie Claire; SHAPIRO, Arnold; HÖGNÄS, Göran. 2005. *Fitting and Forecasting Mortality Rates for Nordic Countries Using the Lee-Carter method*. Actuarial Research Clearing House, 1–21.
- KUČERA, Tomáš. 1998. *Regionální populační prognózy: teorie a praxe prognózování vývoje lidských zdrojů v území*. Disertační práce. Praha: Univerzita Karlova v Praze, Přírodovědecká fakulta.
- LANGHAMROVÁ, Jitka a Tomáš FIALA. 2014. *Předpokládaný vývoj obyvatelstva krajů ČR a jeho důsledky*. Acta Oeconomica Pragensia. Roč. 22, č. 4, 73–96. ISSN 0572-3043.
- LANGHAMROVÁ, Jitka, Tomáš FIALA, Vladimír HULÍK, Martina MISKOLCZI a Eva KAČEROVÁ. 2010. *Prognóza lidského kapitálu obyvatelstva České republiky do roku 2050*. Demografie. Roč. 52 č. 3. str. 181–196. ISSN 0011-8265.



- LEE, Ronald D. 1993. *Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level*. International Journal of Forecasting, 9.2: 187–202.
- LEE, Ronald D. a Lawrence R. CARTER. 1992. *Modeling and Forecasting the Time Series of U. S. Mortality*. Journal of the American Statistical Association, vol. 87, no. 419, pp. 659–671.
- MCCAA, Robert. 1989. *Populate. A microcomputer projection package for aggregative data applied to Norway, 1736–1970*. In: Annales de démographie historique. Persée-Portail des revues scientifiques en SHS. p. 287–298.
- MEDHI, J. 1992. *Statistical methods: an introductory text*. New York: Wiley. ISBN 81-224-0419-7.
- MONTGOMERY, Douglas C., Elizabeth A. PECK a G. Geoffrey VINING. 2012. *Introduction to linear regression analysis*. 5th ed. Hoboken, NJ: Wiley. ISBN 978-0-470-54281-1.
- MYRSKYLÄ, Mikko, Joshua R. GOLDSTEIN a Yen-hsin Alice CHENG. 2013. *New cohort fertility forecasts for the developed world: rises, falls, and reversals*. Population and Development Review, 39.1: 31–56.
- NEUBAUER, Jiří, Marek SEDLAČÍK a Oldřich KRÍŽ. 2016. *Základy statistiky: aplikace v technických a ekonomických oborech*. 2., rozšířené vydání. Praha: Grada. ISBN 978-80-247-5786-5.
- O'DONOVAN, Thomas M. 1983. *Short term forecasting: an introduction to the Box-Jenkins approach*. New York: Wiley. ISBN 0471900133.
- ORD, Keith a Robert FILDES. 2013. *Principles of business forecasting*. Mason, OH: South-Western Cengage Learning. ISBN 978-0-324-31127-3.
- OTYEPKA, Michal, Pavel BANÁŠ a Eva OTYEPKOVÁ. 2013. *Základy zpracování dat*. Univerzita Palackého v Olomouci.
- PAVLÍK, Zdeněk, Jitka RYCHTAŘÍKOVÁ a Alena ŠUBRTOVÁ. 1986. *Základy demografie: celostátní vysokoškolácká příručka pro studenty přírodovědeckých, ekonomických, filozických a lékařských fakult*. Praha: Academia.
- PAVLÍK, Zdeněk a Květa KALIBOVÁ. 2005. *Mnohojazyčný demografický slovník*. 2. vyd., aktual. Praha: Česká Demografická Společnost. ISBN 8023948644.
- PECÁKOVÁ, Iva. 2007. *Logistická regrese s vícekategoriální vysvětlovanou proměnnou*. Acta Oeconomica Pragensia, 1: 86–96.
- PENTREATH, Nick. 2015. *Machine Learning with Spark*. Livery Place Birmingham UK: Packt Publishing. ISBN 978-1-78328-851-9.
- PERISTERA, Paraskevi a Anastasia KOSTAKI. 2007. *Modeling fertility in modern populations*. Demographic Research, 16.6: 141–194.
- PRÁŠKOVÁ, Zuzana. 2016. *Základy náhodných procesů II. Druhé upravené vydání*. Praha: Univerzita Karlova, nakladatelství Karolinum. ISBN 978-80-246-3516-3.

- PRESSAT, Roland. 2008. *Demographic analysis: projections on natality, fertility and replacement*. New Brunswick: AldineTransaction. ISBN 0202361977.
- PRESSAT, Roland. 1995. *Eléments de démographie mathématique*. Paris: Association internationale des démographes de langue française. ISBN 2950935605.
- PRESTON, Samuel H., Patrick HEUVELINE a Michel GUILLOT. 2001. *Demography: measuring and modeling population processes*. Malden, MA: Blackwell Publishers. ISBN 1-55786-214-1.
- ROGERS, Andrei a Luis Javier CASTRO. 1981. *Model migration schedules*. International Institute for Applied Systems Analysis, Laxenburg.
- ROUBAL, Pavel. 2010. *Informatika a výpočetní technika pro střední školy: [kompletní látka pro nižší a vyšší úroveň státní maturity]*. Brno: Computer Press. ISBN 978-80-251-3228-9.
- ROUBÍČEK, Vladimír. 1996. *Základní problémy obecné a ekonomické demografie*. Praha: Vysoká škola ekonomická. ISBN 80-7079-188-8.
- SADALLA, Edward. 2005. *The U.S.–Mexican border environment: dynamics of human–environment interactions*. San Diego, CA: San Diego State University Press. ISBN 0-925613-46-0.
- SHRYOCK, Henry S., Jacob S. SIEGEL, Elizabeth A. LARMON a EDITORS. 1973. *The methods and materials of demography*. Second Printing. Social and Economic Statistics Administrations, US Bureau of the Census.
- SMITH, David P., Nathan KEYFITZ, Kenneth W. WACHTER a Hervé LE BRAS. 2013. *Mathematical demography: selected papers. Second, revised edition*. Berlin: Springer. *Demographic research monographs*. A series of the Max Planck institute for demographic research / Jam. ISBN 9783642358579.
- SWANSON, David A., and Jeff TAYMAN. 2012. *Subnational Population Estimates*. Springer *Series on Demographic Methods and Population Analysis 31*. Springer Science and Business Media. 414 pp. ISBN 978-90-481-8953-3.
- ŠÍDLOVÁ KUNSTOVÁ, Natálie a Luděk ŠÍDLO. 2016. *Perspektivy vývoje kapacit v domovech pro seniory v Jihočeském kraji*. Kontakt. Roč. 18, č. 2, 113–122. ISSN 1212-4117.
- ŠTĚDRŇ, Bohumír, POTŮČEK, Martin, KNÁPEK, Jaroslav, MAZOUCH, Petr a kol. 2012. *Prognostické metody a jejich aplikace*. V Praze: C.H. Beck. Beckova edice ekonomie. ISBN 978-80-7179-174-4.
- ŠTYGLEROVÁ, Terezie, Michaela NĚMEČKOVÁ a Miroslav ŠIMEK. 2010. *Projekce obyvatelstva České republiky 2009–2065*. Demografie. Roč. 52, č. 1, str. 1–14. ISSN 0011-8265.
- VALLIN, Jacques a Graziella CASELLI. 2006. *Relationships between Age-Specific Rates and Synthetic Indicators: Decomposition of a Difference*. In CASELLI, Graziella, Jacques VALLIN a Guillaume J. WUNSCH. 2006. *Demography: analysis and synthesis, Volume I*. London: Elsevier, s. 209–214. ISBN 0-12-765661-8.

- WATKINS, David S. 2002. *Fundamentals of matrix computations*. 2nd ed. New York: Wiley–Interscience. Pure and applied mathematics (John Wiley & Sons: Unnumbered). ISBN 0-471-21394-2.
- WEINSTEIN, Jay A. a Vijayan K. PILLAI. 2016. *Demography: the science of population*. Second edition. ISBN 978-1-4422-3519-9.
- WHELPTON, P. K.. 1928. *The Population of the United States, 1925–1975*. American Journal of Sociology 34, 253–271.
- YAN, Xin a Xiaogang SU. 2009. *Linear regression analysis: theory and computing*. Hackensack, NJ: World Scientific. ISBN 981-283-410-9.
- ZVÁRA, Karel, Patrícia MARTINKOVÁ, Jiří ANDĚL a EDITOR. 2013. *Biomedicínská statistika: Základy statistiky v prostředí R*. 1. vydání. Praha: Nakladatelství Karolinum. ISBN 9788024624471.

## SEZNAM POUŽITÝCH ELEKTRONICKÝCH ZDROJŮ

- ARLT, Josef, Markéta ARLTOVÁ a Eva RUBLÍKOVÁ. 2002. *Analýza ekonomických časových řad s příklady* [online]. Praha: Vysoká škola ekonomická v Praze [cit. 2017-03-16]. Dostupné z: <http://nb.vse.cz/~arltova/vyuka/crsbir02.pdf>.
- BARTKOVÁ, Veronika. 2009. *Věková struktura obyvatelstva Jihočeského kraje a koncepce péče o staré občany* [online]. České Budějovice [cit. 2017-05-20]. Dostupné z: [https://theses.cz/id/uh897i/downloadPraceContent\\_adipIdno\\_10502](https://theses.cz/id/uh897i/downloadPraceContent_adipIdno_10502). Diplomová práce. Jihočeská univerzita. Vedoucí práce Michal Vančura.
- BOOTH, Heather, Leonie TICKLE, John MAINDONALD a Rob J HYNDMAN. 2017. *Package 'demography': Forecasting mortality, fertility, migration and population data* [online]. R package version 1.19 [cit. 2017-04-10]. Dostupné z: <https://cran.r-project.org/web/packages/demography/demography.pdf>.
- CAHOVÁ, Alice. 2012. *Analýza demografického vývoje Jihočeského kraje* [online]. Brno [cit. 2017-05-20]. Dostupné z: [http://is.mendelu.cz/zp/portal\\_zp.pl?prehled=vyhledavani;podrobnosti=45331;zp=33166;download\\_prace=1;lang=cz](http://is.mendelu.cz/zp/portal_zp.pl?prehled=vyhledavani;podrobnosti=45331;zp=33166;download_prace=1;lang=cz). Bakalářská práce. Mendelova univerzita. Vedoucí práce Dana Hübelová.
- ČERMÁKOVÁ, Anna. 2009. *Rizikové faktory karcinomu plic. Plicní karcinom u žen* [online]. Praha [cit. 2017-04-29]. Dostupné z: <https://is.cuni.cz/webapps/zzp/detail/57493>. Univerzita Karlova. Vedoucí práce Jan Kára.
- ČSÚ. 2017a. Český statistický úřad [online]. Praha [cit. 2017-1-17]. Dostupné z: <https://www.czso.cz/>.
- ČSÚ. 2017b. *Vysvětlivky k databázi: Územní změny, počty obyvatel, narození, zemřelí, stěhování* [online]. Praha [cit. 2017-05-04]. Dostupné z: <https://www.czso.cz/csu/czso/vysvetlivky-k-databazi>.
- ČSÚ. 2013e. *Projekce obyvatelstva České republiky do roku 2100* [online]. Praha [cit. 2017-05-19]. Dostupné z: <https://www.czso.cz/csu/czso/projekce-obyvatelstva-ceske-republiky-do-roku-2100-n-fu4s64b8h4>.
- ČSÚ. 2004d. *Populační prognóza ČR do r. 2050 - N* [online]. Praha [cit. 2017-05-19]. Dostupné z: <https://www.czso.cz/csu/czso/populacni-prognoza-cr-do-r2050-n-g9kah2fe2x>.

- HYNDMAN, Rob J. 2017. *Package 'forecast': Forecasting functions for time series and linear models* [online]. R package version 8.0 [cit. 2017-04-10]. Dostupné z: <https://cran.r-project.org/web/packages/forecast/forecast.pdf>.
- HYNDMAN, Rob J a George ATHANASOPOULOS. 2017. *ARIMA modelling in R* [online]. OTexts [cit. 2017-04-28]. Dostupné z: <https://www.otexts.org/fpp/8/7>.
- JIHOČESKÝ KRAJ. 2013. *Program rozvoje Jihočeského 2014 – 2020: Příloha č. 3 socioekonomický profil Jihočeského* [online]. KP projekt s.r.o. a Krajský úřad Jihočeského kraje, OREG [cit. 2017-05-21]. Dostupné z: [http://www.kraj-jihocesky.cz/file.php?par\[id\\_r\]=96472&par\[view\]=0](http://www.kraj-jihocesky.cz/file.php?par[id_r]=96472&par[view]=0)
- MALTHUS, Thomas. 1798. *An Essay on the Principle of Population: An Essay on the Principle of Population, as it Affects the Future Improvement of Society with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers*. [online]. London: Printed for J. Johnson, in St. Paul's Church-Yard [cit. 2017-05-19]. Dostupné z: <http://www.esp.org/books/malthus/population/malthus.pdf>.
- MARTINKOVÁ, Lucie. 2012. *Lineární modely časových řad a jejich aplikace na vybraných ekonomických problémech* [online]. Olomouc [cit. 2017-04-12]. Dostupné z: <http://theses.cz/id/yky203/00164587-954239442.pdf>. Diplomová práce. Univerzita Palackého v Olomouci. Vedoucí práce Ondřej Vencálek.
- MATHCENTRE. 2013. *Variance and standard deviation (grouped data)* [online]. [cit. 2017-02-24]. Dostupné z: [http://www.lboro.ac.uk/media/wwwlboroacuk/content/mlsc/downloads/var\\_stand\\_deviat\\_group.pdf](http://www.lboro.ac.uk/media/wwwlboroacuk/content/mlsc/downloads/var_stand_deviat_group.pdf).
- MATYAŠOVSKÁ, Jana. 2014. *Analýza volebného správania na Slovensku* [online]. Bratislava [cit. 2017-05-12]. Dostupné z: <http://www.iam.fmph.uniba.sk/efm/bakalarky/2014/matyasovska/bakalarka.pdf>. Bakalářská práce. Univerzita Komenského v Bratislave. Vedoucí práce Martin Niepel.
- MYŠÁKOVÁ, Gabriela. 2011. *Metody analýzy sezónnosti demografických jevů* [online]. Praha [cit. 2017-04-11]. Dostupné z: <https://is.cuni.cz/webapps/zzp/detail/76675>. Univerzita Karlova. Vedoucí práce Klára Hulíková.
- POHOŘÁLKOVÁ, Petra. 2007. *Zdravotní pojištění v prostředí stárnutí populace* [online]. Praha [cit. 2017-05-26]. Dostupné z: <https://is.cuni.cz/webapps/zzp/detail/44344>. Bakalářská práce. Univerzita Karlova. Vedoucí práce Petr Bohumský.
- POLÁŠKOVÁ, Michaela. 2013. *Konstrukce úmrtnostních tabulek a jejich využití v pojišťovnictví* [online]. Olomouc [cit. 2017-02-11]. Dostupné z: <http://theses.cz/id/f9vio4/00171312-255451601.pdf>. Bakalářská práce. Univerzita Palackého v Olomouci. Vedoucí práce Ondřej Vencálek.
- RAYMER, James, Arkadiusz WIŚNIEWSKI, Peter WF SMITH a Jakub BIJAK. 2012. *Applying and Extending the Lee-Carter Model to Forecast Migration* [online]. ESRC Centre for Population Change, University of Southampton [cit. 2017-04-10]. Dostupné z: [https://iussp.org/sites/default/files/event\\_call\\_for\\_papers/leecarter\\_migration.pdf](https://iussp.org/sites/default/files/event_call_for_papers/leecarter_migration.pdf)

- SKŘIVANOVÁ, Zuzana. 2016. *Stochastické modelování úmrtnosti pro více populací* [online]. Praha [cit. 2017-03-13]. Dostupné z: <https://is.cuni.cz/webapps/zzp/detail/172273>. Diplomová práce. Univerzita Karlova v Praze. Vedoucí práce Lucie Mazurová.
- ŠTÁDLEROVÁ, Kateřina. 2013. *Projekce úmrtnosti podle příčin úmrtí* [online]. Praha [cit. 2017-04-05]. Dostupné z: <https://is.cuni.cz/webapps/zzp/detail/114055>. Univerzita Karlova. Vedoucí práce Jan Kořístka.
- UNITED NATIONS. 2013a. *MORTPAK FOR WINDOWS: Mortpak Manual (Version 4.3)* [online]. United Nations, New York [cit. 2016-11-23]. Dostupné z: [http://www.un.org/en/development/desa/population/publications/pdf/mortality/mortpak\\_manual.pdf](http://www.un.org/en/development/desa/population/publications/pdf/mortality/mortpak_manual.pdf).
- ZADRAŽILOVÁ, Jana. 2016. *Mateřství v pozdním věku aneb fenomén dnešní doby* [online]. Praha [cit. 2017-05-01]. Dostupné z: <https://vskp.vse.cz/id/1284420>. Bakalářská práce. Vysoká škola ekonomická. Vedoucí práce Martina Miskolczi.

## **SEZNAM POUŽITÝCH SOFTWAREVÝCH PRODUKTŮ**

- BOOTH, Heather, Leonie TICKLE, John MAINDONALD a Rob J HYNDMAN. 2014. Demography: Forecasting mortality, fertility, migration and population data [online]. R package version 1.18 [cit. 2016-12-14]. Dostupné z: <https://CRAN.R-project.org/package=demography>.
- BURCIN, Boris, Klára Hulíková TESÁRKOVÁ a Komárek DAVID. 2012. DeRaS: software tool for modelling mortality intensities and life table construction [online]. Charles university in Prague, Prague [cit. 2017-04-13]. Dostupné z: <http://deras.natur.cuni.cz>.
- HYNDMAN, Rob J. 2016. Forecast: Forecasting functions for time series and linear models [online]. R package version 7.2 [cit. 2016-12-18]. Dostupné z: <http://github.com/robjhyndman/forecast>.
- MATLAB. 2012b. The MathWorks, Inc., Natick, Massachusetts, United States.
- R CORE TEAM. 2016. R: A language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna, Austria [cit. 2016-09-03]. Dostupné z: <https://www.R-project.org/>.
- SAS INSTITUTE INC. 2017. Base SAS 9.4 Procedures Guide. Cary, NC, USA.
- UNITED NATIONS. 2013b. MORTPAK for Windows, Version 4.3. New York, USA.

## SEZNAM POUŽITÝCH DATOVÝCH ZDROJŮ

### ÚMRTNOST

- ČSÚ. 2016a. *Úmrtnostní tabulky za ČR, regiony soudržnosti a kraje – 2014 – 2015. Podrobné úmrtnostní tabulky za kraje ČR v letech 2014 – 2015: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-regiony-soudrznosti-a-kraje>
- ČSÚ. 2016b. *Úmrtnostní tabulky za ČR, regiony soudržnosti a kraje – 2014 – 2015. Podrobné úmrtnostní tabulky za kraje ČR v letech 2014 – 2015: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-regiony-soudrznosti-a-kraje>
- ČSÚ. 2015a. *Úmrtnostní tabulky za ČR, regiony soudržnosti a kraje – 2013 – 2014. Podrobné úmrtnostní tabulky za kraje ČR v letech 2013 – 2014: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-regiony-soudrznosti-a-kraje-2013-az-2014>
- ČSÚ. 2015b. *Úmrtnostní tabulky za ČR, regiony soudržnosti a kraje – 2013 – 2014. Podrobné úmrtnostní tabulky za kraje ČR v letech 2013 – 2014: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-regiony-soudrznosti-a-kraje-2013-az-2014>
- ČSÚ. 2014a. *Úmrtnostní tabulky za ČR, regiony soudržnosti a kraje – 2012 až 2013. Podrobné úmrtnostní tabulky za kraje ČR v letech 2012 – 2013: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-regiony-soudrznosti-a-kraje-2012-az-2013-opauclw7zk>
- ČSÚ. 2014b. *Úmrtnostní tabulky za ČR, regiony soudržnosti a kraje – 2012 až 2013. Podrobné úmrtnostní tabulky za kraje ČR v letech 2012 – 2013: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-regiony-soudrznosti-a-kraje-2012-az-2013-opauclw7zk>
- ČSÚ. 2013a. *Úmrtnostní tabulky za ČR 2012, oblasti a kraje – 2011 až 2012. Podrobné úmrtnostní tabulky za kraje ČR v letech 2011 – 2012: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-2012-oblasti-a-kraje-2011-az-2012-0o9j8toftl>



- ČSÚ. 2013b. *Úmrtnostní tabulky za ČR 2012, oblasti a kraje – 2011 až 2012. Podrobné úmrtnostní tabulky za kraje ČR v letech 2011 – 2012: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-2012-oblasti-a-kraje-2011-az-2012-0o9j8toft1>
- ČSÚ. 2012a. *Úmrtnostní tabulky za ČR, oblasti a kraje – 2010, 2011. Podrobné úmrtnostní tabulky za kraje ČR v letech 2010 – 2011: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-oblasti-a-kraje-2010-2011-5zvr1xygig>
- ČSÚ. 2012b. *Úmrtnostní tabulky za ČR, oblasti a kraje – 2010, 2011. Podrobné úmrtnostní tabulky za kraje ČR v letech 2010 – 2011: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-oblasti-a-kraje-2010-2011-5zvr1xygig>
- ČSÚ. 2011a. *Úmrtnostní tabulky za ČR, oblasti a kraje – 2010. Podrobné úmrtnostní tabulky za kraje ČR v letech 2009 – 2010: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-oblasti-a-kraje-2010-q4oxyosuq2>
- ČSÚ. 2011b. *Úmrtnostní tabulky za ČR, oblasti a kraje – 2010. Podrobné úmrtnostní tabulky za kraje ČR v letech 2009 – 2010: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-oblasti-a-kraje-2010-q4oxyosuq2>
- ČSÚ. 2010a. *Úmrtnostní tabulky za ČR, oblasti a kraje – 2008 – 2009. Podrobné úmrtnostní tabulky za kraje ČR v letech 2008 – 2009: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-oblasti-a-kraje-2008-2009-9wq3wyazpu>
- ČSÚ. 2010b. *Úmrtnostní tabulky za ČR, oblasti a kraje – 2008 – 2009. Podrobné úmrtnostní tabulky za kraje ČR v letech 2008 – 2009: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-08-05]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-oblasti-a-kraje-2008-2009-9wq3wyazpu>
- ČSÚ. 2009a. *Úmrtnostní tabulky za ČR a kraje – 2009. Podrobné úmrtnostní tabulky za kraje ČR v letech 2007–2008: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-09-19]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2009-4zcg1m6zyr>
- ČSÚ. 2009b. *Úmrtnostní tabulky za ČR a kraje – 2009. Podrobné úmrtnostní tabulky za kraje ČR v letech 2007–2008: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-09-19]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2009-4zcg1m6zyr>
- ČSÚ. 2008a. *Úmrtnostní tabulky za ČR a kraje – 2006 – 2007. Podrobné úmrtnostní tabulky za kraje ČR v letech 2006 – 2007: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-09-19]. Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2006-2007-5taldl9e0q>

- ČSÚ. 2008b. *Úmrtnostní tabulky za ČR a kraje – 2006 – 2007. Podrobné úmrtnostní tabulky za kraje ČR v letech 2006 – 2007: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2006-2007-5taldl9e0q>
- ČSÚ. 2007a. *Úmrtnostní tabulky za ČR a kraje – 2005 – 2006. Podrobné úmrtnostní tabulky za kraje ČR v letech 2005 – 2006: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2005-2006-17t453zjb1>
- ČSÚ. 2007b. *Úmrtnostní tabulky za ČR a kraje – 2005 – 2006. Podrobné úmrtnostní tabulky za kraje ČR v letech 2005 – 2006: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2005-2006-17t453zjb1>
- ČSÚ. 2006a. *Úmrtnostní tabulky za ČR a kraje – 2004 – 2005. Podrobné úmrtnostní tabulky za kraje ČR v letech 2004 – 2005: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2004-2005-llpickene0>
- ČSÚ. 2006b. *Úmrtnostní tabulky za ČR a kraje – 2004 – 2005. Podrobné úmrtnostní tabulky za kraje ČR v letech 2004 – 2005: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2004-2005-llpickene0>
- ČSÚ. 2005a. *Úmrtnostní tabulky za ČR a kraje – 2003 – 2004. Podrobné úmrtnostní tabulky za kraje ČR v letech 2003 – 2004: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-09-20].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2003-2004-597pq3avf8>
- ČSÚ. 2005b. *Úmrtnostní tabulky za ČR a kraje – 2003 – 2004. Podrobné úmrtnostní tabulky za kraje ČR v letech 2003 – 2004: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-09-20].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2003-2004-597pq3avf8>
- ČSÚ. 2004a. *Úmrtnostní tabulky za ČR a kraje – 2002 – 2003. Podrobné úmrtnostní tabulky za kraje ČR v letech 2002 – 2003: Tab. Jihočeský kraj, muži* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2002-2003-qq02ejp0dt>
- ČSÚ. 2004b. *Úmrtnostní tabulky za ČR a kraje – 2002 – 2003. Podrobné úmrtnostní tabulky za kraje ČR v letech 2002 – 2003: Tab. Jihočeský kraj, ženy* [online]. [cit. 2016-09-19].  
Dostupné z: <https://www.czso.cz/csu/czso/umrtnostni-tabulky-za-cr-a-kraje-2002-2003-qq02ejp0dt>

## PLODNOST

ČSÚ. 2016c. *Demografické ročenky (pramenná díla) 2010 – 2015: Tab. Míry plodnosti podle jednotek věku, věkových skupin žen, typu urbanizace a krajů (na 1000 žen)* [online]. [cit. 2016-08-12]. Dostupné z: [https://www.czso.cz/csu/czso/casova\\_rada\\_demografie](https://www.czso.cz/csu/czso/casova_rada_demografie)

ČSÚ. 2010c. *Demografické ročenky (pramenná díla) 2009 – 2008: Tab. Míry plodnosti podle jednotek věku, věkových skupin žen, typu urbanizace a krajů (na 1000 žen)* [online]. [cit. 2016-08-12]. Dostupné z: [https://www.czso.cz/csu/czso/casova\\_rada\\_demografie\\_2009\\_1990](https://www.czso.cz/csu/czso/casova_rada_demografie_2009_1990)

ČSÚ. 2008c. *Demografické ročenky (pramenná díla) 2007 – 2000: Tab. Míry plodnosti podle jednotek věku, věkových skupin žen a krajů (na 1000 žen)* [online]. [cit. 2016-09-19]. Dostupné z: [https://www.czso.cz/csu/czso/casova\\_rada\\_demografie\\_2009\\_1990](https://www.czso.cz/csu/czso/casova_rada_demografie_2009_1990)

## VĚKOVÁ STRUKTURA OBYVATELSTVA

ČSÚ. 2016d. *Věkové složení obyvatelstva – 2015: Tab. Věkové složení mužů k 1. 1. 2015* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva>

ČSÚ. 2016e. *Věkové složení obyvatelstva – 2015: Tab. Věkové složení žen k 1. 1. 2015* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva>

ČSÚ. 2015c. *Věkové složení obyvatelstva – 2014: Tab. Věkové složení mužů k 1. 1. 2014* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2014>

ČSÚ. 2015d. *Věkové složení obyvatelstva – 2014: Tab. Věkové složení žen k 1. 1. 2014* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2014>

ČSÚ. 2014c. *Věkové složení obyvatelstva – 2013: Tab. Věkové složení mužů k 1. 1. 2013* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2013-3sqlgzbrnf>

ČSÚ. 2014d. *Věkové složení obyvatelstva – 2013: Tab. Věkové složení žen k 1. 1. 2013* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2013-3sqlgzbrnf>

ČSÚ. 2013c. *Věkové složení obyvatelstva – 2012: Tab. Věkové složení mužů k 1. 1. 2012* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2012-jf4obhd6b8>

ČSÚ. 2013d. *Věkové složení obyvatelstva – 2012: Tab. Věkové složení žen k 1. 1. 2012* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2012-jf4obhd6b8>

ČSÚ. 2012c. *Věkové složení obyvatelstva – 2011: Tab. Věkové složení mužů k 1. 1. 2011* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2011-kikapm7yay>

ČSÚ. 2012d. *Věkové složení obyvatelstva – 2011: Tab. Věkové složení žen k 1. 1. 2011* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2011-kikapm7yay>

- ČSÚ. 2011c. *Věkové složení obyvatelstva – 2010: Tab. Věkové složení mužů k 1. 1. 2010* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2010-bj74aj1afw>
- ČSÚ. 2011d. *Věkové složení obyvatelstva – 2010: Tab. Věkové složení žen k 1. 1. 2010* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2010-bj74aj1afw>
- ČSÚ. 2010d. *Věkové složení obyvatelstva – 2009: Tab. Věkové složení mužů k 1. 1. 2009* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2009-9oav0o3wtg>
- ČSÚ. 2010e. *Věkové složení obyvatelstva – 2009: Tab. Věkové složení žen k 1. 1. 2009* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2009-9oav0o3wtg>
- ČSÚ. 2009c. *Věkové složení obyvatelstva – 2008: Tab. Věkové složení mužů v ČR k 1. 1. 2008* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2008-bsgh7zhdwg>
- ČSÚ. 2009d. *Věkové složení obyvatelstva – 2008: Tab. Věkové složení žen v ČR k 1. 1. 2008* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2008-bsgh7zhdwg>
- ČSÚ. 2008d. *Věkové složení obyvatelstva – 2007: Tab. Věkové složení obyvatel ČR podle pohlaví a věku k 1. 1. 2007* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-2007-j4ysr151mt>
- ČSÚ. 2007c. *Věkové složení obyvatelstva České republiky – 2006. Kraje: věkové složení obyvatel podle pohlaví a věku v roce 2006: Tab. Jihočeský kraj* [online]. [cit. 2016-07-23]. Dostupné z: [https://www.czso.cz/csu/czso/4003-07-v\\_roce\\_2006-50n](https://www.czso.cz/csu/czso/4003-07-v_roce_2006-50n)
- ČSÚ. 2006c. *Věkové složení obyvatelstva České republiky – 2005. Kraje: věkové složení obyvatel podle pohlaví a věku v roce 2005: Tab. Jihočeský kraj* [online]. [cit. 2016-07-23]. Dostupné z: [https://www.czso.cz/csu/czso/4003-06-v\\_roce\\_2005-50n](https://www.czso.cz/csu/czso/4003-06-v_roce_2005-50n)
- ČSÚ. 2005c. *Věkové složení obyvatelstva ČR – 2004: Tab. Věkové složení obyvatel v Jihočeském kraji podle pohlaví a věku v roce 2004* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-cr-2004-xyr87e3dj8>
- ČSÚ. 2004c. *Věkové složení obyvatelstva České republiky – 2003: Tab. Věkové složení obyvatel v Jihočeském kraji podle pohlaví a věku v roce 2003* [online]. [cit. 2016-07-23]. Dostupné z: <https://www.czso.cz/csu/czso/vekove-slozeni-obyvatelstva-ceske-republiky-2003-t9gwaekzgi>

## OSTATNÍ DATOVÉ ZDROJE

- BURCIN, Boris, Zdeněk ČERMÁK, Tomáš KUČERA a Luděk ŠÍDLLO. 2014. *Prognóza vývoje počtu obyvatel v krajích České republiky do roku 2065: Východiska, předpoklady a základní výsledky prognózy*. Výzkumný ústav vodohospodářský T. G. Masaryka, veřejná výzkumná instituce.

- ČSÚ. 2014e. *Projekce obyvatelstva v krajích ČR – do roku 2050* [online]. Praha [cit. 2017-05-19]. Dostupné z: <https://www.czso.cz/csu/czso/projekce-obyvatelstva-v-krajich-cr-do-roku-2050-ua08v25hx9>
- ČSÚ. 2016f. *Obyvatelstvo: Tab. Časové řady* [online]. [cit. 2016-09-01]. Dostupné z: <https://www.czso.cz/csu/xc/obyvatelstvo-xc>